Are priority sites for shorebird and seabird conservation adequately represented in Kenya's Marine Protected Areas?

Martin Mwema Musangu

March, 2012

Course Title:	Geo-Information Science and Earth Observation for Environmental Modelling and Management (GEM)
Level:	Master of Science (Msc)
Course Duration:	September 2010 - March 2012
Consortium partners:	
	Lund University (Sweden)
	International Institute for Geo-Information Science and Earth Observation (ITC) (The Netherlands)
GEM thesis number:	

Are priority sites for shorebird and seabird conservation adequately represented in Kenya's Marine Protected Areas?

by

Martin Mwema Musangu

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation for Environmental Modelling and Management

Thesis Assessment Board

Chairman: Prof. Dr. A.K. (Andrew) Skidmore
External Examiner: Prof. Petter Pilesjö
First Supervisor : Drs. Raymond Nijmeijer (ITC)
Second Supervisor : Dr. A.G. (Bert) Toxopeus (ITC)

ITC
UNIVERSITY OF TWENTE.
FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Abstract

Marine protected areas (MPAs) in Kenya were initially established for objectives other than conservation of seabirds and shorebirds. Therefore, it is unclear whether high priority areas for conservation of these birds are adequately represented within Kenya's MPA system. To investigate this, a gap analysis was carried out using Marxan. Since fine resolution species distribution data required by Marxan was not available from existing atlases, Species Distribution Modelling (SDM) using Maxent (Maximum Entropy Model) was carried out to generate distribution maps for input. Results show that high priority areas for conservation of seabirds and shorebirds are not adequately represented within Kenya's MPAs. Only about a quarter of these priority sites are found within the current MPA system. Five major gaps were identified by Marxan. Of these, three are already recognized as Important Bird Areas (IBAs) i.e. Tana River Delta, Kisite area and Sabaki River Mouth. The other two, Lamu Archipelago and Ngomeni area, are not IBAs and therefore there is a need for further research to determine whether they meet the thresholds to become IBAs. Results also showed that land cover was the most important factor determining shorebird and seabird distribution. The environmental and human factors studied did not significantly influence shorebirds and seabirds differently with the exception of distance from rivers. Observations from the field showed that seabirds can shift breeding locations from more protected areas to less protected or unprotected areas. To ensure the long-term survival of shorebirds and seabirds in Kenya, the areas identified as gaps need to be protected under the MPA network.

Key words: Shorebirds and seabirds, marine protected areas, species distribution modelling, gap analysis

Acknowledgements

My sincere gratitude goes to my funders, the World Wide Fund for Nature (WWF), who supported my Msc including the fieldwork through the Russell E. Train education for nature scholarship. This opportunity has been the fulfilment of a long cherished ambition. I further extend my thanks to Ms. Stephanie Eisenman at WWF for the smooth administration of the scholarship.

I wish to thank my supervisors Drs. Raymond Nijmeijer (ITC), Dr. Bert Toxopeus (ITC) and Dr. Erustus Kanga (Kenya Wildlife Service-KWS) for guidance throughout the process and for always being available whenever I needed advice. *Asante sana.* In addition, Prof. Andrew Skidmore, Dr. Alexei Voinov and Ir. Louise van Leeuwen offered helpful suggestions during my mid-term and proposal defence that greatly improved my study. Many thanks to Prof. Petter Pilesjö and Mr. Ulrik Mårtensson in Sweden who assisted me in securing a Swedish visa even when I had almost given up hope.

I owe special thanks to Colin Jackson for advice and guidance in the field especially on identification of seabirds, sampling and a gentle introduction to the tides. Thanks also to the staff at Arocha Kenya for a warm reception during my stay there. Special thanks as well to Josphine Mutiso (KWS) whose valuable advice on logistics helped me to easily find my way to all the sampling areas. I thank the following who have assisted me in carrying out this study in one way or another: Zeno Wijtten (GVI), John Musina (NMK) and Dr. Mohammed (KWS). Many thanks to Elijah Machira for hosting me at the coast.

I thank my Mum, Jane Machira, for her love and unwavering belief in me. I am also indebted to the Waheire family for their support and hospitality throughout my studies. My heartfelt thanks go to Jane Ndungu for advice and encouragement throughout the thesis writing period. In equal measure, I thank the Kenyan community at ITC for their friendship and companionship and in particular to Salome Wabuyele for all the wonderful get-togethers.

Last but certainly not least, I wish to thank all my GEM classmates for their companionship and friendship during the entire time we have been together. You have been like a family to me.

Table of Contents

List of figures

List of tables

1. Introduction

Marine Protected Areas (MPAs) in Kenya cover an area of about 706 km^2 (UNEP, 1998) which is equivalent to about 8.1% of the continental shelf (Wells *et al.*, 2007). The first MPAs were established in the early 1960's primarily to protect coral reefs while second generation MPAs were established to play additional roles such as tourism, protection of fish stocks and research (Francis *et al.*, 2002). Although Kenya's MPAs were not created with the explicit aim of seabird and shorebird protection, they play an important role in the protection of these species with several Important Bird Areas (IBAs) for seabirds and shorebirds being located within MPAs (Wells*, et al.*, 2007). However, it is unclear to what degree these species and their habitats are represented within the MPA system. In order for MPAs to effectively play their role in protecting seabirds and shorebirds, the entire suit of species and their critical habitats have to be represented within the MPAs in sufficient quantities to ensure their long term persistence (Margules & Pressey, 2000).

Gap analysis is a useful method of finding out which important species or habitats are not represented in the MPA system. At its simplest form, it involves overlaying species distribution maps with protected area maps to identify species ranges outside protected area boundaries i.e. gaps (Rodríguez *et al.*, 2007). There are several software products available with the common ones being Gap Analysis Program (Scott *et al.*, 1993) and Marxan (Watts *et al.*, 2009).

Marxan is a tool for reserve design that has been used for, among other things, determining where gaps exist in Protected Area (PA) networks as well as determining the efficiency of the PA network in meeting conservation targets (Ardron *et al.*, 2010). It is also used to identify which priority sites are unrepresented or inadequately represented within the protected area system (Esselman & Allan, 2011). Since its release in 1999, it has become popular and has been used for gap analysis and reserve planning by organizations such as the World Wide Fund for Nature (WWF) and The Nature Conservancy (TNC) (Ardron, *et al.*, 2010). Given that it is increasingly becoming difficult to set aside land for protected areas due to competing community interests and activities, any process that seeks to identify priority sites with a view to recommending them for conservation needs to take account the socio-economic costs. Marxan does this by using a simulated annealing algorithm to identify areas that meet all species conservation targets at the lowest cost. To achieve this, Marxan minimises the following cost factor (Ball & Possingham, 2000).

$$\underbrace{\sum_{Sites}\text{Cost}}_{1} + \underbrace{BLM\sum_{Sites}\text{Boundary}}_{2} + \underbrace{\sum_{ConValue} CFPF \times \text{Penalty}}_{3} + \underbrace{\text{Cost Threshold Penalty}(t)}_{4}$$

Where:

1. Total cost of the reserve network
2. The total reserve boundary length, multiplied by a modifier
3. The penalty for not adequately representing conservation features
4. The penalty for exceeding a preset cost threshold

The first cost refers to planning unit costs which may be measures of socio-economic costs or undesirable factors within the site. Marxan minimises this by avoiding selecting units with high costs such as those near roads or in urban areas. Boundary refers to cost associated with the boundary length and the Boundary Length Modifier (BLM) determines importance of the costs associated with the boundary. Marxan seeks to minimize boundary length so that you have small compact areas (Ball & Possingham, 2000). Marxan has several strengths especially in selection of priority sites for conservation. It ensures connectivity between selected areas which promotes recruitment thus making it possible to sustain local populations. One can also decide on minimum reserve size to ensure that only sustainable reserves are selected. In addition, it is possible to select only areas with the least opportunity cost of establishment (Stewart *et al.*, 2003).

Effective gap analysis requires accurate information on species distribution. In many areas, especially in developing countries, this information may not be available due to the inadequate resources for research and monitoring in terms of funds and technical capacity (Esselman & Allan, 2011). To work around this problem, Species Distribution Models (SDM's) are now commonly used to predict species occurrence where data is lacking or limited (Rodríguez, *et al.*, 2007). SDM's use statistical relationships between species distribution and habitat conditions to predict the probability of a species occurring in an area (Guisan & Zimmermann, 2000). This allows for the extrapolation of incomplete information to yield distribution maps that take into account unsurveyed areas (Rondinini *et al.*, 2006). This is an advantage in Kenya where information on birds is biased towards popular birding destinations and is often lacking in insecure areas. SDM's may also produce finer resolution data compared to atlas maps which have coarser resolutions (50 by 50 km in Kenya's case) and may include areas unsuitable for certain species (Lawler *et al.*, 2010). There are several SDM's in use including the Habitat

Suitability Index (HSI), Mahalanobis distance, ecological niche models like Genetic Algorithm for Rule-Set Production (GARP) (Johnson & Gillingham, 2005), Generalized Linear Models (GLM), Random Forest (RF) and maximum entropy (Maxent) (Marini *et al.*, 2010). Maxent is advantageous in distribution modelling because it requires presence only data and smaller sample sizes (Esselman & Allan, 2011).

1.1 Problem Statement

The Kenyan coast is an important wintering ground for migratory seabirds and shorebirds. The islands and islets are also important breeding areas for several seabird species (UNEP, 1998). Three of the five IBAs along the Kenyan coastline are not protected under the MPA system (Wells*, et al.*, 2007) and therefore the species, habitats and processes within these areas are susceptible to threats that may have a negative impact on them. The biggest imminent threats include plans to convert over 200,000 hectares of the Tana Delta for biofuels and sugar cane production (Birdlife International, 2011b). There is also a proposal to build a multi-billion dollar port in Lamu to cater for increased transport demand due to the recent discovery of oil in Uganda and the independence of South Sudan given that both countries are land-locked (The East African, 2009). Lamu is near several important seabird breeding sites in Kiunga (Figure 1). Hence, it is important to come up with an objective process to identify high priority sites for conservation of seabirds and shorebirds and assess their degree of representation within the MPA network. Areas or species that are excluded or under-represented can be prioritized and recommended for conservation action either through expansion of the MPA network or through other avenues such as community conservation or land purchases. This study combines species distribution modelling and gap analysis to determine whether seabirds and shorebirds are adequately represented in Kenya's MPAs.

1.2 Research Objectives

Overall objective

To identify priority sites for conservation of shorebirds and seabirds and determine whether these sites are adequately represented within Kenya's MPAs.

Specific objectives

i) To determine shorebird and seabird distribution ranges and to identify factors influencing their distribution.

ii) To identify priority sites for the conservation of shorebirds and seabirds.

iii) To determine whether the identified priority sites are adequately represented within Kenya's MPAs.

1.3 Research Questions

1.3.2 Overall research question

Are priority sites for conservation of shorebirds and seabirds adequately represented within Kenya's MPAs?

1.3.2 Specific research questions

i) What environmental and human factors influence the distribution of shorebirds and seabirds? Do these factors influence seabirds and shorebirds differently?

ii) Which are the priority sites for conservation of shorebirds and seabirds? Do these sites correspond to priority sites identified by experts through the IBA process?

iii) Which priority sites for seabirds and shorebirds have been excluded from the MPA network? Do MPAs and IBAs contain larger areas of priority sites within their boundaries than outside?

1.4 Research Hypotheses

i) There is a significant difference in the influence of factors between seabirds and shorebirds.

ii) The total area of the identified priority sites within IBAs is significantly smaller than the total area outside the IBA.

iii) The total area of identified priority sites within MPAs is significantly smaller than the total area outside the MPA.

2. Materials and Methods

2.1 Study Area

The study area covers the entire Kenyan coastline which stretches over a distance of about 600 km from the Tanzanian border in the South (Longitude 4° 40' S) to the Somali border in the North (Longitude 4° 40' S) (Figure 1) (Government of Kenya, 2009). It extends 2 km landward and 2 km seaward.

The weather in this area is influenced by the monsoons which blow from the Northeast between December and March and from the Southeast from May to October. As a result, the long rainy season occurs between March and May while the short rainy season occurs between October and December. Rainfall ranges from 500-1600 mm with rainfall increasing as you move towards the South. Mean temperatures range between 24 ^{o}C and 30 ^{o}C (Government of Kenya, 2009). Seabirds normally breed during the South-eastern monsoon season (June to October) when strong winds and rough seas make access to the islands difficult thereby reducing incidences of egg collection (Birdlife International, 2011a).

The study area encompasses a variety of habitats including seagrass meadows, coral reefs, creeks, sandy beaches, salt pans, offshore islands, mangroves and estuaries (Francis*, et al.*, 2002). The estuaries support congregations of shorebirds, many of them migratory. Mangrove swamps on the other hand, provide breeding areas and shelter for fish and these in turn serve as food for seabirds. Over 50 species of coastal shorebirds and seabirds have been recorded in this region with over half of them being Palaearctic migrants (Zimmerman *et al.*, 1999).

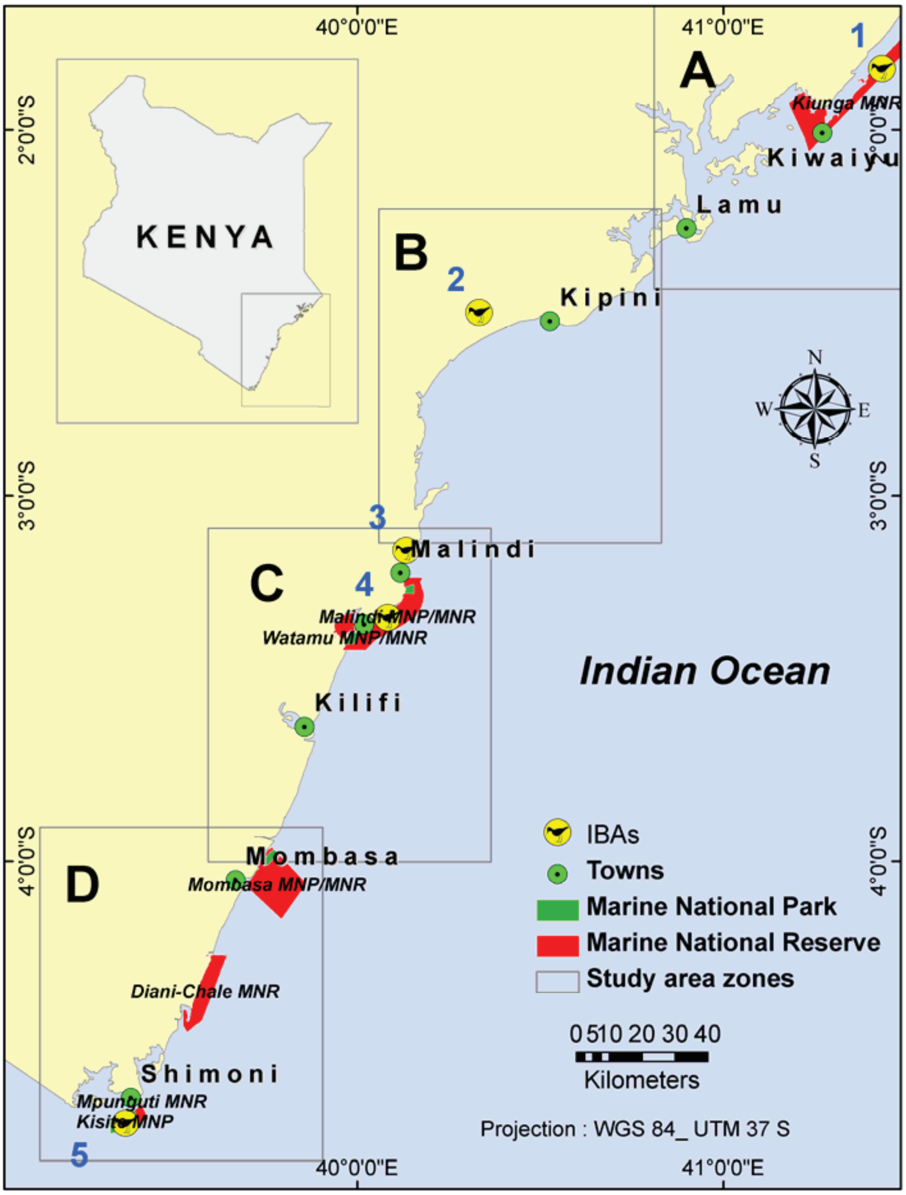


Figure 1: Study area map

There are five designated IBAs within the study area containing either significant congregations or breeding sites of shorebirds and seabirds (Bennun & Njoroge, 1999). These are: Kiunga (1), Tana River Delta

(2), Sabaki River Mouth (3), Watamu area (4, which incorporates Mida Creek, Whale Island and Malindi) and Kisite Island (5) (Figure 1). Of these Sabaki River Mouth and Tana River Delta are not represented within the MPA system but there are proposals to create the Tana Delta wetland reserve to address this gap (UNEP, 1998). Currently there are ten marine protected areas covering an area of about 706 km^2. Four are Marine National Parks (MNP) i.e. Malindi MNP, Watamu MNP, Mombasa MNP and Kisite MNP. The other six are Marine National Reserves (MNR) i.e. Malindi MNR, Watamu MNR, Mombasa MNR, Kiunga MNR, Diani-Chale MNR and Mpunguti MNR (Figure 1). Reserves generally have a lower degree of protection with traditional fishing being allowed within their boundaries (UNEP, 1998).

2.2 Methods

The procedures carried out during the study are summarized in Figure 2.

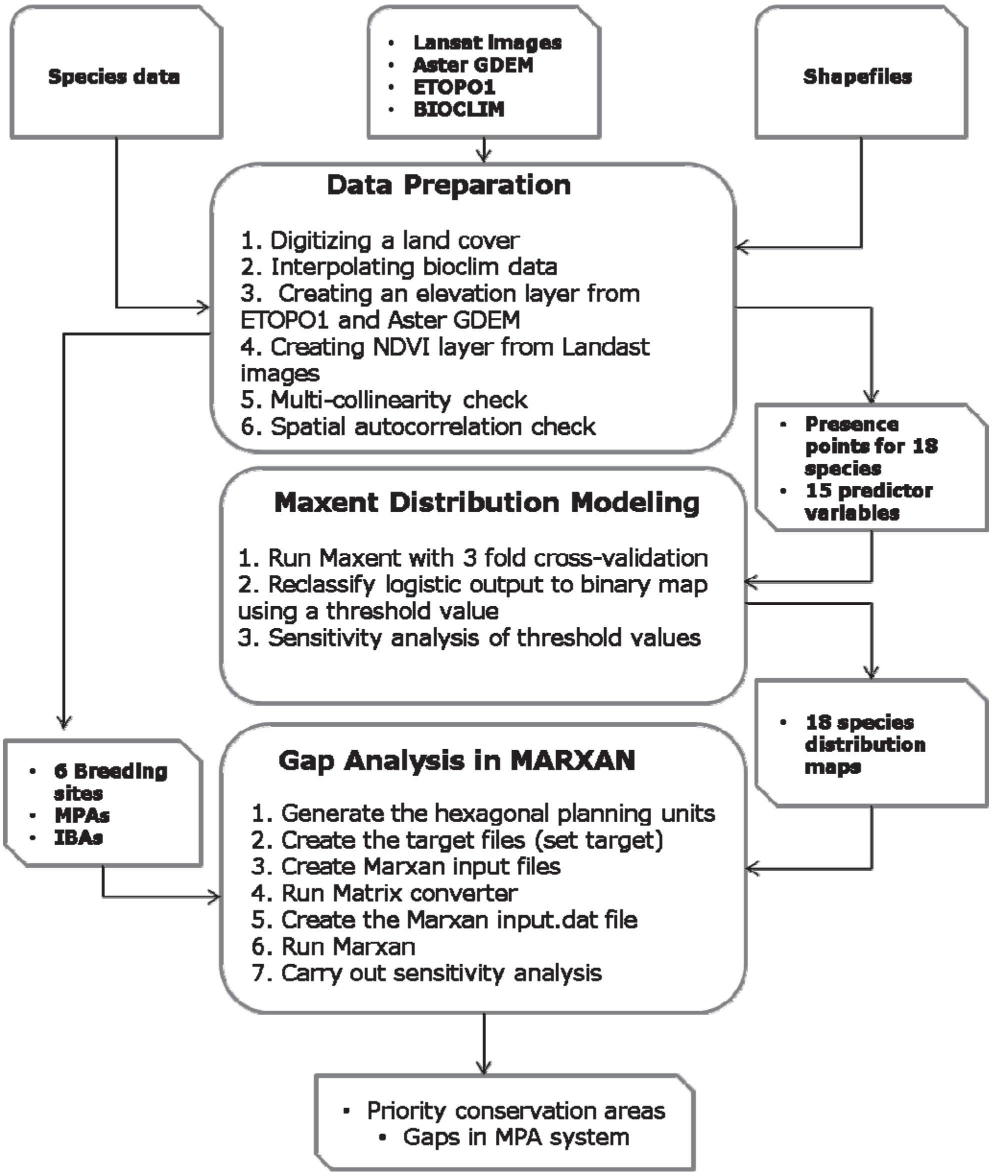


Figure 2: Method flow chart

2.2.1 Fieldwork

Given that the exact locations (GPS points) available for species occurrence were not sufficient, fieldwork was carried out to collect additional field observations points. The fieldwork was carried out for one month between 12 September and 12 October 2011 during the breeding season of most seabirds. It involved collecting both primary and secondary data. Primary data was collected from Shimoni near the Tanzanian border to North of Malindi at Ngomeni (Figure 1), a coastline length of about 300 km. Areas North of Malindi were avoided due to insecurity from Somali pirates. For those areas, secondary data was collected from institutions and individuals involved in monitoring birds such as the Ornithology department of the National Museums of Kenya and Arocha Kenya (a conservation-based NGO).

Primary data collection was done using a stratified systematic sampling strategy. The 300 km coast where primary data collection occurred was stratified into 6 sections of 50 km in order to account for the variation in climatic and habitat conditions as you move from the South towards the North. In each 50 km section, at least 7 transects, each 2 km long were made. In total 42 transects were established corresponding to 42 sampling points. However, effort was made to ensure that all designated IBAs within that section were sampled. Transects were sampled both by boat and on foot depending on accessibility and availability of boats. In each transect, the following was done:

i) Recording of GPS points at the beginning and end of each transect, as well as points where there are congregations of shorebirds and seabirds.

ii) Identification and counting of all seabird and shorebird species.

ii) Recording of habitat characteristics and taking of habitat photos.

2.2.2 Data sources and preparation

Species data was obtained from fieldwork and secondary sources. A major secondary source was the results of a survey carried out in 1995 (Nasirwa *et al.*, 1995) while additional records were gotten from Kenya Birdfinder, an internet based bird sightings database (Appendix 1). Of the 50 species of shorebirds and seabirds that have been recorded at the Kenyan coast, 23 species were selected for the gap analysis (Table 1). Species that were excluded were vagrant species (i.e. recorded less than 10 times in Kenya), pelagic seabirds with no breeding sites in Kenya and species that were not restricted to the Kenyan coast. Priority was given to threatened species, resident breeding birds and migratory species. Since data was collected from a variety of sources, there was need to check for accuracy of the data especially for secondary data. This involved checking the accuracy of the coordinates provided, names of locations and species abundance. Secondary and primary data were then combined and harmonized by ensuring the use of similar names for species and sites and removal of duplicate records. The species presence points from both primary and secondary data were then checked for spatial autocorrelation in ArcGIS 9.3. This involved extracting data values of predictor variables at each presence point and then testing for spatial autocorrelation using Morans I index (Sokal & Oden, 1978). Moran's I values can range from -1 (indicating strong negative autocorrelation), 0 (indicating a random pattern) to 1 (indicating a strong positive autocorrelation). Values of Z-scores greater than 1.96 or less than -1.96 show significant spatial autocorrelation (at $\alpha=0.05$) (Carnes & Ogneva-Himmelberger, 2011).

Most of the factors had low spatial autocorrelation with the exception of some bioclim variables (Appendix 3).

Efforts to find a detailed land cover /habitat map showing the various habitats that are important for seabirds and shorebirds were unsuccessful. Particularly lacking were accurate maps of the sandy beaches and intertidal reef flats as well as seabird breeding islands and coral platforms. It was therefore necessary to create one through on-screen digitization. A combination of Google earth images, existing shapefiles and PDF habitat maps from the Kensea project (Tychsen, 2006) (Appendix 1) were used. Google earth images were chosen because they are free and have very good resolution for most areas of the Kenyan coast since the images are obtained from GeoEye. To enable use of Google earth photos, zooming was done until the desired level (i.e. 10 m altitude) and then saved as JPEG pictures. This yielded over 200 images. These were stitched into groups using Adobe Photoshop for easier geo-referencing. The resulting groups were then geo-referenced using roads and country outline shapefiles. The PDF maps were converted to TIFF and then also geo-referenced. A geodatabase was then created so as to enable the defining of topology rules to ensure there was no overlap of polygons or gaps. Digitizing was carried out and then topology checks were conducted. The resulting map was exported as a shapefile and then rasterized at a resolution of 30m using habitat codes (Appendix 2).

The Normalized Difference Vegetation Index (NDVI) layer also had to be created since the available NDVI layers excluded the intertidal and oceanic zones. Landsat 5 TM images with a resolution of 30m (Appendix 1) were used to make the NDVI layer. Images with low cloud cover (less than 10%) acquired between 2009 and 2011 were used. In addition, only images acquired when the intertidal areas were exposed (i.e. at low tide) were used. Since Landsat 5 acquires images over the equator at around 10 GMT (13:00 East Africa Time), historical tide timetables (Appendix 1) were used to check whether it was low or high tide at that time on the particular day the image was acquired. The ERDAS ATCOR module was then used for atmospheric correction, haze removal and calculation of NDVI.

The elevation layer presented a problem because some of the species presence points used in modelling fell on land and others were located in the intertidal zones. The landward side had high resolution Aster GDEM (30m resolution) images available (Appendix 1) but there was no corresponding high resolution bathymetry images. The highest resolution bathymetry images found were the ETOPO1 images which have a resolution of 1 arc minute (approx. 1.85 km resolution) (Appendix 1). Therefore, the ETOPO1 images were resampled to 30m and then mosaicked with the Aster GDEM (with the Aster GDEM being given priority where there was overlap).

Bioclimatic (bioclim) data from WorldClim (Appendix 1) was also used and the data comprised of 18 bioclimatic layers mainly, temperature, precipitation and derivations from the two. The bioclim data was also not available for the intertidal and oceanic areas since the layers are provided when already clipped using the country's coastline. Interpolation using inverse distance weighted (IDW) method was therefore done in order to extend the coverage of these layers into the intertidal areas. Interpolation using IDW retains the original values in cells that already had values. However, interpolation creates floating point values. To convert them to integer values and round-off the values in each pixel, the following formula was used in the ArcGIS raster calculator: output_raster = Int ([input_raster] +0.5).

Several layers were created showing the Euclidian distance of each cell/pixel to different human and environmental factors. The factors investigated were distances from:

- Rivers
- Roads
- Urban areas and settlements
- Fish landings
- Hotels
- Marine Protected Areas
- Mangroves

All the input layers were converted to the same projection (UTM zone 37S) and then clipped using the study area extent. They were then converted to ASCII grid files ready for input into Maxent.

2.2.3 Selection of variables and checking for multi-collinearity

A multi-collinearity check was conducted to see whether there was high correlation among the continuous predictor variables (Appendix 3). This was done by doing a regression of one factor against the remaining factors. The resulting coefficient of determination was converted into Variance Inflation Factors (VIF). VIF is calculated using the formula (O'brien, 2007):

$$VIF = \frac{1}{1-R^2}$$

Where R^2 is the coefficient of determination

Variables with high inflation factors (more than 10) (O'brien, 2007) were considered collinear and were removed unless they were considered to be ecologically important for the species (Appendix 3).

2.2.4 Maxent Modelling Procedure

Maxent requires species data and predictor variables data (i.e. environmental/human factor layers) as input. The predictor variables need to be in ASCII format and the species presence points need to be in csv format. The ASCII raster layers also need to have the same extent, geographic projection, cell size and number of rows and columns. It is also necessary to specify which predictor variables are continuous and which are categorical.

In this study, Maxent modelling was done using 18 species of seabirds and shorebirds (Table 1) and 15 predictor variables (Table 2). Of the predictor variables, all were continuous variables except land cover which was a categorical variable.

Table 1: Species considered in the gap analysis

Species	**Scientific Name**	**No of Pts**	**Status**	**Type**
Species modelled with Maxent				
Grey Plover	*Pluvialis squatarola*	43	M	Shorebird
Whimbrel	*Numenius phaeopus*	43	M	Shorebird
Sooty Gull	*Larus hemprichii*	35	M	Seabird
Terek Sandpiper	*Xenus cinereus*	31	M	Shorebird
Great Egret	*Casmerodius albus*	29	R	Shorebird
Greater Sandplover	*Charadrius leschenaultii*	29	M	Shorebird
Lesser Crested Tern	*Sterna bengalensis*	24	M	Seabird
Gull-billed Tern	*Gelochelidon nilotica*	23	M	Seabird
Ruddy Turnstone	*Arenaria interpres*	21	M	Shorebird
Sanderling	*Calidris alba*	21	M	Shorebird
Common Tern	*Sterna hirundo*	18	M	Seabird
Saunders's Tern	*Sterna saundersi*	18	M	Seabird
Greater Crested Tern	*Sterna bergii*	15	M	Seabird
Lesser Sandplover	*Charadrius mongolus*	15	M	Shorebird
White-fronted Plover	*Charadrius marginatus*	14	R	Shorebird
Caspian Tern	*Sterna caspia*	12	M	Seabird
Crab Plover	*Dromas ardeola*	12	M	Shorebird
Eurasian Curlew	*Numenius arquata*	11	M	Shorebird
Seabirds with breeding location data				
Bridled Tern	*Sterna anaethetus*		R	Seabird
Sooty Tern	*Sterna fuscata*		R	Seabird
Roseate Tern	*Sterna dougalii*		R	Seabird
White-cheeked Tern	*Sterna repressa*		R	Seabird
Brown Noddy	*Anous stolidus*		R	Seabird
Sooty Gull	*Larus hemprichii*		R	Seabird

M-Migratory R-Resident

Table 2: Predictor variables used in modelling with Maxent

Factor code	Factor name	Units
Bio13	Precipitation of Wettest Month	mm
Bio15	Precipitation Seasonality	mm
Bio18	Precipitation of Warmest Quarter	mm
Bio3	Isothermality (BIO2/BIO7) (* 100)	°C * 10
Bio9	Mean Temperature of Driest Quarter	°C * 10
Dist_hotel	Distance from Hotels	meters
Cover/cov	Land cover	meters
Dist_fish	Distance from fish landings	meters
Dist_mang	Distance from Mangroves	meters
Dist_mpa	Distance from Marine protected areas	meters
Dist_river	Distance from rivers	meters
Dist_road	Distance from roads	meters
Dist_urban	Distance from urban areas	meters
Elevation/elev	Elevation	meters
ndvi	NDVI	meters

The model was run using a 3-fold cross–validation with options for output of response curves and a jacknife test. Default values were used for all other parameters. Cross-validation enables all the data to be used both for training and for validation. This is possible because the data is split into folds (e.g. 3) and one fold is used for validation and the rest for creating the model. It is therefore convenient for use with small number of samples (Phillips, 2005).

A logistic output map of presence probability of between 0 and 1 was generated by Maxent. To determine presence or absence, this map was reclassified using the equal training sensitivity and specificity logistic threshold (Cantor *et al.*, 1999) (Table 4). To determine what effect threshold would have on the high priority areas selected by Marxan, a maximum sensitivity plus specificity threshold was also used for comparison. This threshold has considerably lower values

(Table 4). Both thresholds have previously been found to give good results (Liu *et al.*, 2005). A sensitivity analysis of the thresholds was carried out by increasing or decreasing the value and accessing the impact on the distribution. A threshold was considered too high if the original presence points were omitted from the distribution range or too low if the distribution showed the species occurring in habitats that it is not known to occur based on expert knowledge. The sensitivity analysis showed that the equal training sensitivity and specificity logistic threshold was the most appropriate for use in this study.

Validation of the model was done using the area under Receiver Operating Characteristic (ROC) curve (AUC) which is produced as an output by Maxent (Table 4). Training AUC shows the fit to the training data.

2.2.5 Marxan Gap Analysis Procedure

Gap analysis was carried out on the 18 species modelled by Maxent as well as on breeding locations of 6 seabird species (Table 1). To prepare files for input into Marxan, the Protected Area Tools (PAT tools) was used. This is a free decision support software that was developed by The Nature Conservancy (TNC) and it is compatible with ArcGIS 9.1 to 10 (Appendix 1). Marxan Tools, a module within PAT tools, was used to generate all the files required for running of Marxan.

In order for Marxan to run, it requires the following files: a planning unit file (pu.dat), a species file (spec.dat), a planning unit versus species matrix file (puvsp.dat) and an input file (input.dat). Optional files to run are a boundary file (bound.dat) and a block definition file (block.dat). For this study, all files except the block.dat were created using the Marxan input generator in Marxan Tools.

The planning unit file (pu.dat) contains three fields: planning unit id, cost and status. To enable the planning unit file to be created, planning units were first generated. These are grids or cells over the study area extent on which analysis takes place. Each planning unit has a specific id. For this study, a 100 ha (1 km^2) hexagonal planning unit grid was generated using the Hexgen tool in Marxan tools. The planning unit cost was generated using the Environmental Risk Surfaces (ERS) module in PAT tools. This generates a cost layer based on mapped risks. Settlements, roads and fish landings were used as inputs. The cost generated by ERS was added to the cost of a planning unit based on area to get the final cost of a planning unit. Marxan tries to minimise cost by selecting the cheapest planning units that meet the targets.

The species file (spec.dat) contains the following fields: species, proportion/target (prop) and species penalty factor (spf). The species file gives the id of each species that was used in the analysis and the proportion of the distribution range or breeding area that should be conserved for each species. For this study, the target was set at 100% (1.0) for all breeding sites and 40% (0.4) for the distribution ranges based on literature (Delavenne *et al.*, 2011; IUCN, 2003) and species requirements. The species penalty factor is a penalty that is assigned if a target is not met. A high spf (10000) was used to ensure that all targets are met. Formatting of the input shapefiles (distribution ranges and breeding sites) needed to be done to enable their use by the Marxan input generator which uses them to prepare a spec.dat file. This was done using the Target Prep in Marxan tools which dissolves repeating rows and adds a field for specifying the proportion to be conserved (target) and a target id field.

The planning unit versus species matrix file (puvsp.dat) has three fields: the species (which gives the id of the species), the pu (which is the planning unit id) and the amount (which shows the amount of the distribution range or breeding site within each planning unit). The boundary file (bound.dat) contains information of boundary costs of adjacent units.

The inedit.exe file was used to create the input file (input.dat). When inedit.exe is run it gives options where one can specify the location of the input files (spec.dat, pu.dat, puvsp.dat), specify input and output directories, specify types of output files desired and set different parameters (Table 3). Zonae Cogito was then used to run Marxan and to display the outputs. It was also used for calibration and sensitivity analysis of the various parameters. ArcGIS was used to format and to produce maps from the Marxan output.

Table 3: Parameter values used in running Marxan

Parameter	**Value**
Boundary length modifier	1.0
Repeat Runs	100
Input file type	New Free form
Run options	simulated annealing- normal iterative improvement
number of iterations	1000000
temperature decrease	10000
Optional input files	boundary length file (block.dat)
Screen Output	General output

Defaults were used for all other parameters

2.3 Analysis

2.3.1 Identifying Environmental and human factors influencing distribution of seabirds and shorebirds

Results from a Jacknife test using Maxent (i.e. jacknife values of regularized training gain) were used to ascertain which factors were the most important for the distribution of seabirds and shorebirds. When a jacknife option is selected, Maxent creates several models. First, one variable is excluded and the remaining variables are used to create the model. This is done for all the variables in turn. In the second model, only one variable is used to create the model. In the final model, all variables are used. The gain of the model using both training and test data is given as outputs (Phillips, 2005).

To test whether the factors influence seabirds and shorebird differently, a non-parametric test (Mann-Whitney U-test) was used. Jacknife values of training model made using one variable at a time (Appendix 3 and 4) were used for the analysis where the test checked whether there was a significant difference in the training gains of the different factors between seabirds and shorebirds. The Mann-Whitney U-test was chosen due to the small and unequal sample sizes which were likely not to meet the assumption of normality (seabirds, n=7, shorebirds, n=11).

2.3.2 Identifying priority sites for conservation of shorebirds and seabirds

To identify priority sites for conservation of seabirds and shorebirds, Marxan was first run without restrictions (i.e. without locking in any MPAs or IBAs). This first run used Maxent output maps generated using equal sensitivity and specificity threshold. This was done to see if Marxan would select areas that are already protected or designated as IBAs. This would indicate how well the current MPAs and IBAs represent seabirds and shorebirds. Numbers of selected cells that fall in MPAs and IBAs were then determined and chi-square test was done to determine if there is a significant difference between:

i) Selected planning units inside and outside MPAs.

ii) Selected planning units inside and outside IBAs.

iii) Selected planning units inside and outside MPAs and IBAs combined.

The procedure was repeated using the Maximum sensitivity plus specificity threshold (Table 4) and the two outputs were compared to investigate the effect of using different thresholds.

2.3.3 Identifying priority sites for conservation of seabirds and shorebirds that are excluded from the MPA network (Gaps)

To identify gaps given the current protected area system, Marxan was run with the MPA layer locked in so that planning units within MPAs are given first priority during selection. Only when it is not possible to fulfil the given targets from within the MPA planning units does Marxan select from outside the MPA. Marxan was also run with the MPA and IBA layers locked in to find out what would be the gaps if all current IBAs would be conserved in the future.

3. Results

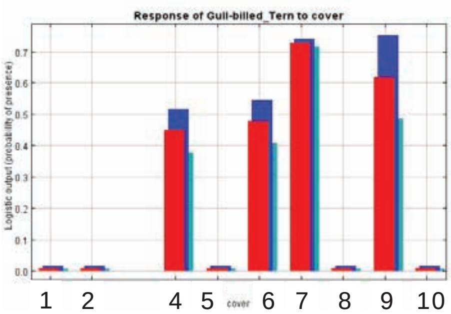
3.1 Maxent model Validation

All the 18 species modelled had high test AUC values with the Caspian Tern (with 12 modelled points) having the lowest test AUC value and the Crab Plover (with 12 modelled points) having the highest (Table 4).

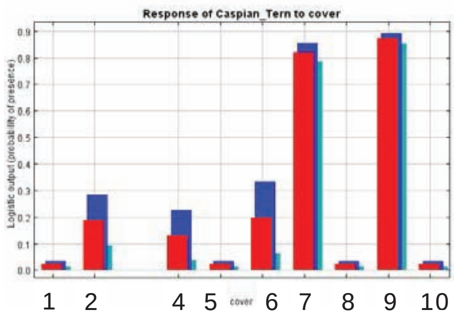
Table 4: Summary results of model validation

Species (number of modelled points)	Test AUC	AUC Standard Deviation	Equal training sensitivity and specificity	Maximum test sensitivity plus specificity logistic threshold
Caspian Tern (12)	0.92	0.06	0.1	0.34
Common Tern (18)	0.99	0.01	0.2	0.03
Crab Plover (12)	0.99	0.002	0.2	0.18
Eurasian Curlew (11)	0.98	0.02	0.1	0.04
Greater Crested Tern (15)	0.94	0.03	0.1	0.26
Great Egret (29)	0.96	0.02	0.2	0.10
Greater Sandplover (29)	0.94	0.03	0.1	0.12
Grey Plover (43)	0.98	0.01	0.1	0.07
Gull-billed Tern (23)	0.94	0.03	0.1	0.03
Lesser Crested Tern (24)	0.97	0.02	0.1	0.03
Lesser Sandplover (15)	0.99	0.01	0.2	0.27
Ruddy Turnstone (21)	0.98	0.01	0.1	0.05
Sanderling (21)	0.97	0.02	0.1	0.14
Saunders's Tern (18)	0.98	0.01	0.1	0.02
Sooty Gull (35)	0.97	0.02	0.1	0.04
Terek Sandpiper (31)	0.97	0.02	0.1	0.07
Whimbrel(43)	0.98	0.01	0.2	0.12
White-fronted Plover (14)	0.94	0.03	0.1	0.05

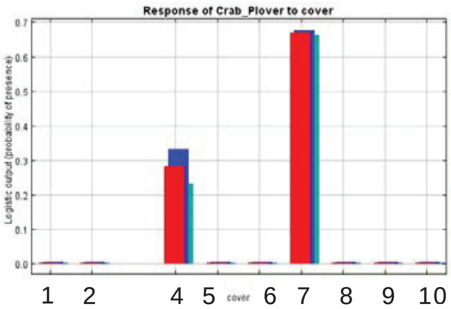
Species that were habitat specific (i.e. occurred in fewer habitat types) such as the Crab Plover, had higher AUC values compared to those that occurred on more habitat types. (Fig. 5, C).



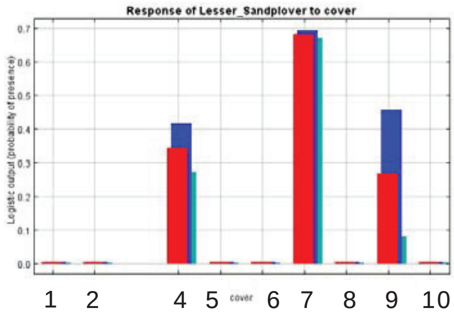
A) Gull-billed Tern (23 points, 0.93 AUC)



B) Caspian Tern (12 points, 0.92 AUC)



C) Crab Plover (12 points, AUC 0.99)



D) Lesser Sandplover (15 points, AUC 0.99)

Figure 3: Comparison of AUC values with number of habitat a species occurs

3.2 Environmental and human factors influencing distribution of seabirds and shorebirds

Land cover was the most important factor in determining the distribution of seabirds and shorebirds along the Kenyan coastline. Results of the jacknife test of variable importance from the Sooty Gull (Figure 4, Table 5) show that when the model is run with all other factors except cover, the training gain of the model reduces considerably from around 3.9 to 3.1. This indicates that land cover provides the most information (i.e. that is not present in other variables) in determining distribution of the species. Influence of other factors was minimal since training gain reduced only marginally when the specific factor was omitted. This same trend was observed for all the species modelled (Table 5 and 6).

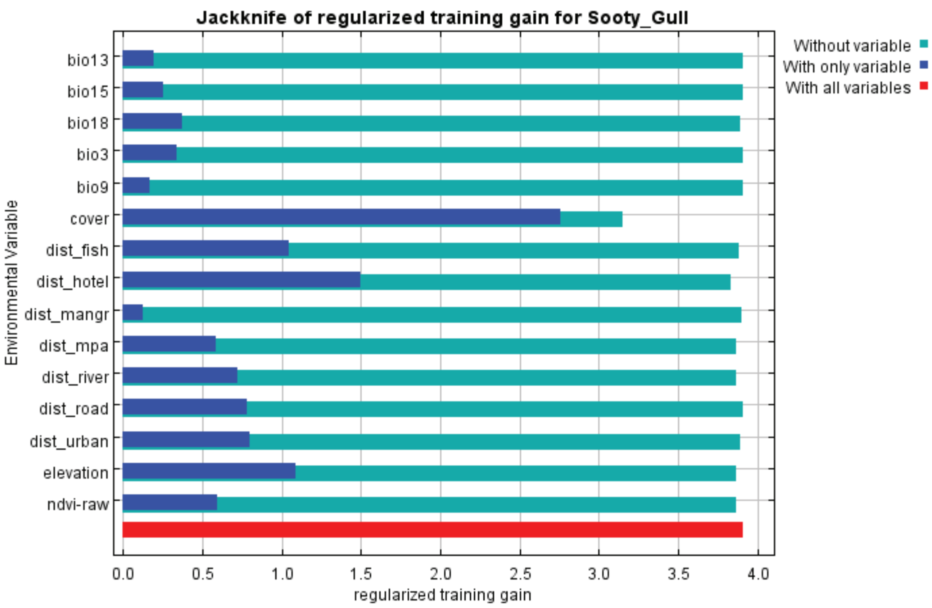


Figure 4: Jacknife values of training gain for Sooty Gull

Table 5: Summary of training gain values for seabirds with one factor excluded

Training gain without variable	**Caspian tern**	**Common tern**	**Greater Crested Tern**	**Gull-billed Tern**	**Lesser Crested Tern**	**Saunders's Tern**	**Sooty Gull**
bio13	4.3	4.6	4.7	3.3	3.5	4.3	3.9
bio15	4.3	4.6	4.7	3.3	3.6	4.4	3.9
bio18	4.2	4.6	4.7	3.2	3.5	4.3	3.9
bio3	4.3	4.6	4.7	3.3	3.6	4.4	3.9
bio9	4.3	4.6	4.7	3.3	3.6	4.3	3.9
cover	3.8	4.0	4.0	2.4	2.9	3.6	3.1
dist_fish	4.3	4.6	4.7	3.3	3.5	4.3	3.9
dist_hotel	4.2	4.4	4.7	3.3	3.4	4.3	3.8
dist_mangr	4.3	4.6	4.7	3.3	3.6	4.4	3.9
dist_mpa	4.2	4.5	4.6	3.2	3.6	4.4	3.9
dist_river	4.2	4.6	4.6	3.2	3.5	4.3	3.9
dist_road	4.3	4.6	4.7	3.3	3.6	4.4	3.9
dist_urban	4.2	4.6	4.6	3.2	3.5	4.3	3.9
elevation	4.3	4.6	4.7	3.2	3.5	4.3	3.9
ndvi-raw	4.3	4.6	4.6	3.2	3.5	4.3	3.9

When the effect of each variable in isolation was considered for the Sooty Gull, the trend was the same with cover contributing most to the training gain (Figure 4, Appendix 5). The trend was also the same for all the other species. For four species (i.e. Caspian Tern, Common Tern, Ruddy Turnstone and Sanderling), distance to hotel was also important (Appendix 4, 5).

For both seabirds and shorebirds, the most preferred land cover types were sand or mud (7) intertidal reef flats (4) and saltworks (9)

(Figure 3: A, B). Some species (mostly seabirds e.g. Caspian Tern and Gull-billed Tern) also had preference for the seasonal wetland cover type (6) (Figure 3: A, B). For all other land cover types, there was little or no preference.

Table 6: Training gain values for shorebirds with one factor excluded

Excluded Variable	**Euraian curlew**	**Grey Plover**	**Lesser Sandplover**	**Ruddy Turnstone**	**Sanderling**	**Terek Sandpiper**	**Whimbrel**	**White-fronted Plover**	**Crab Plover**	**Great Egret**	**Greater Sandplover**
bio13	4.3	3.4	4.7	4.5	4.4	3.6	3.3	4.5	4.0	3.5	3.7
bio15	4.3	3.4	4.7	4.5	4.4	3.6	3.3	4.5	4.0	3.5	3.7
bio18	4.3	3.4	4.6	4.5	4.3	3.6	3.3	4.5	4.0	3.4	3.7
bio3	4.3	3.4	4.7	4.5	4.4	3.6	3.3	4.5	4.0	3.5	3.7
bio9	4.3	3.4	4.7	4.5	4.4	3.6	3.3	4.5	4.0	3.5	3.7
cover	3.6	2.6	4.1	4.1	3.8	2.4	2.4	3.5	2.6	2.6	2.6
dist_fish	4.3	3.4	4.7	4.5	4.3	3.5	3.3	4.5	4.0	3.4	3.7
dist_hotel	4.3	3.4	4.6	4.5	4.3	3.5	3.3	4.5	3.9	3.4	3.6
dist_mangr	4.3	3.4	4.7	4.5	4.3	3.6	3.3	4.5	4.0	3.5	3.6
dist_mpa	4.2	3.4	4.6	4.5	4.3	3.5	3.3	4.5	3.9	3.4	3.7
dist_river	4.3	3.4	4.7	4.5	4.4	3.5	3.3	4.4	4.0	3.5	3.7
dist_road	4.3	3.4	4.6	4.5	4.3	3.6	3.3	4.5	4.0	3.5	3.7
dist_urban	4.3	3.4	4.7	4.5	4.3	3.6	3.3	4.5	3.9	3.5	3.6
elevation	4.3	3.4	4.7	4.5	4.3	3.6	3.3	4.5	4.0	3.5	3.7
ndvi	4.2	3.4	4.5	4.4	4.2	3.6	3.2	4.5	4.0	3.5	3.6

The effect of some of the factors on the probability of occurrence can be found in response curves (Figure 5). The probability of occurrence is greatest about 5 km from roads and urban areas (Figure 5: H, I).

When distance from MPAs and distance from Mangroves are considered, some species displayed a double peak (Figure 5: F) and others displayed a single peak (Figure 5: G).

3.2.2 Do factors affect seabirds and shorebirds differently?

When the model was run with only one factor at a time, all the factors studied, with the exception of distance from a river, did not affect seabirds and shorebirds differently (Table 7). Distance from a river had had a significantly higher training gain ($\alpha = 0.05$) on seabirds than shorebirds (Table 7) indicating a greater influence of this factor on seabirds than shorebirds. This influence can be seen in the response curves where seabirds (Appendix 7A) generally had higher probabilities of occurring near rivers than shorebirds (Appendix 7B).

Table 7: Mann-Whitney U test results summary

Factor	U	DF 1	DF 2	P-value
Bio13	58.000	11	7	0.085
Bio15	58.000	11	7	0.085
Bio18	58.000	11	7	0.085
Bio3	51.000	11	7	0.285
Bio9	58.000	11	7	0.085
Dist_hotel	53.000	11	7	0.211
Cover	50.000	11	7	0.328
Dist_fish	42.000	11	7	0.791
Dist_mang	53.000	11	7	0.211
Dist_mpa	39.000	11	7	1.000
Dist_river	65.000	11	7	0.015*
Dist_road	53.000	11	7	0.211
Dist_urban	52.000	11	7	0.246
elevation	51.000	11	7	0.285
ndvi	40.000	11	7	0.930

DF 1 (shorebirds), DF 2 (seabirds)

Nevertheless, for some factors, response curves showed some seabird species having different trends from those of shorebirds. For example, when land cover was considered, seabirds such as the Lesser Crested Tern (Figure 5: B), Gull-billed Tern (Figure 3: A) and Caspian Tern (Figure 3: B) had a high probability of being found in seasonal wetlands (code 6) compared to the shorebirds. When distance from mangroves was considered most seabirds (i.e. four of the seven seabird species) displayed a double peak in occurrence probability as can be seen with the Great Crested Tern (Figure 5: D) while most shorebirds had a single peak closer to the mangroves similar to the Whimbrel (Figure 5: E).

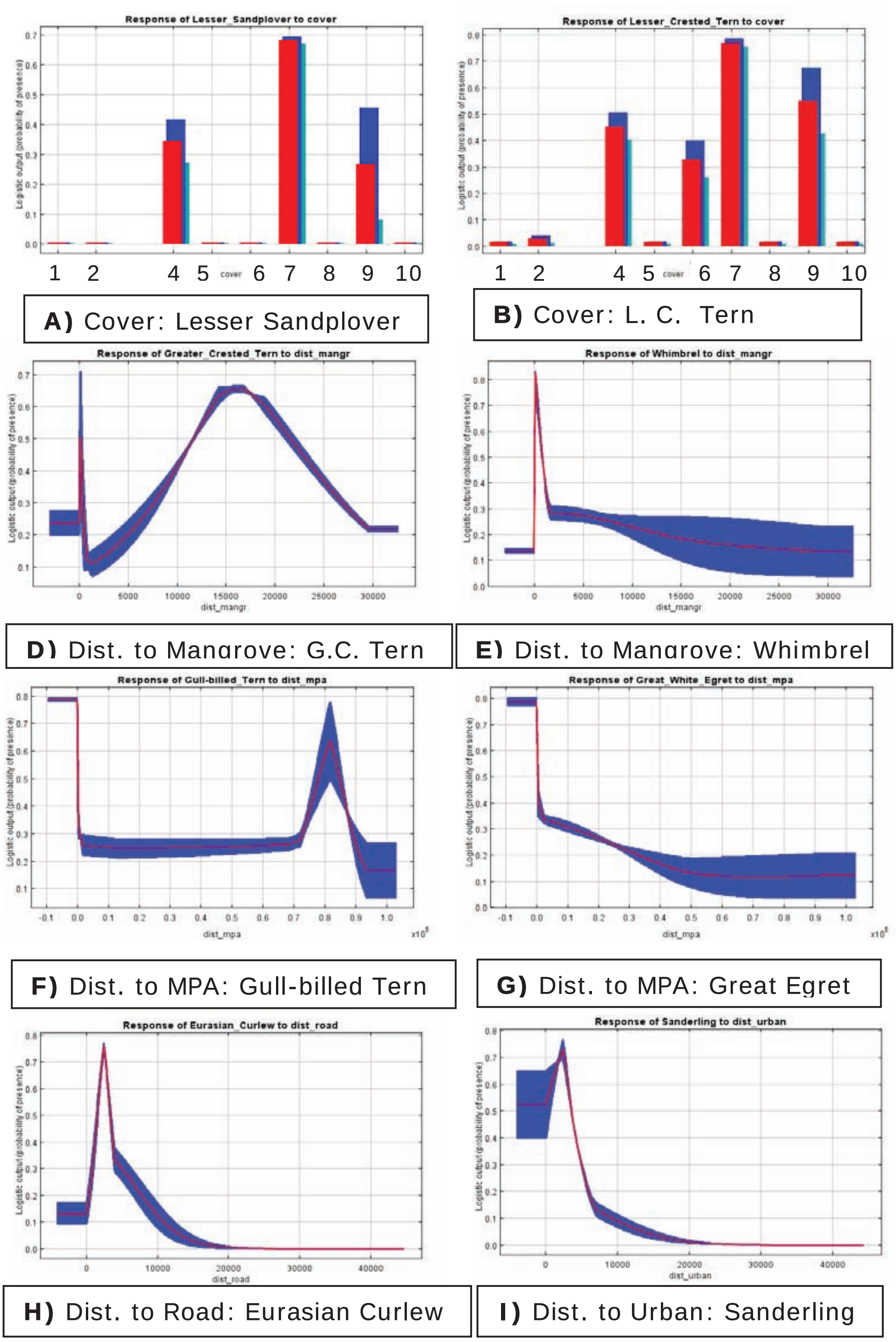


Figure 5: Response curves showing effects of predictor variables

3.3 Priority sites for conservation of shorebirds and seabirds

When Marxan was run without restrictions, several large clusters were identified especially around:

- Kiunga area (north of Kiwaiyu - Figure 6: A)
- Lamu Archipelago (west of Lamu - Figure 6: A, B)
- Tana River Delta (east of Kipini - Figure 6: B)
- Ngomeni area(east of Kipini - Figure 6: B)
- Sabaki River Mouth (south of Mambrui - Figure 6: C)
- Watamu area (around watamu and Mida - Figure 6: C),
- Mombasa area (including north of Shimo la Tewa - Figure 6: C, D)
- Diani area (including north of Gazi - Figure 6: D)
- Kisite area (south east of Shimoni - Figure 6: D)

Of these areas identified, Mombasa area, Diani area, Kiunga area and Watamu area are already within MPAs. When equal training specificity and sensitivity threshold was used, the proportion of planning units selected by Marxan outside MPAs was significantly larger than those inside MPAs (X^2 = 105.5, df = 1, $p < 0.05$) (Table 8). The proportion of selected planning units was also significantly larger outside IBAs (X^2 = 14.5, df = 1, $p < 0.05$) (table 8). However, the area of selected planning units inside and outside MPAs and IBAs combined was not statistically significant (X^2 = 0.45, df = 1, $p = 0.05$) (Table 8). However, the IBAs had better representation of high priority areas (41.4%) compared to MPAs (27%). A combination of the two had the best representation of 48.5%. (Table 8)

When the maximum sensitivity plus specificity threshold was used, the same areas were identified as high priority areas for seabirds and shorebirds (Appendix 6) but the proportion of high priority areas outside of MPAs and IBAs was slightly higher (Table 8). The area covered by the identified priority areas was however almost double that identified using equal sensitivity and specificity threshold. The proportion of selected planning units was significantly larger outside MPAs ($X^2 = 212.49$, df = 1, $p < 0.05$), IBAs ($X^2 = 39.16$, df = 1, $p < 0.05$) and MPAs/IBAs combined ($X^2 = 11.97$, df = 1, $p < 0.05$) (Table 8).

Table 8: Summary of selected planning units in MPAs and IBAs using different thresholds

	Area outside (km^2)	**Area inside (km^2)**
Equal training specificity and sensitivity threshold		
MPA only	363 (73 %)	134 (27 %)
IBA only	291 (58.6%)	206 (41.4%)
IBA and MPA	256 (51.5%)	241 (48.5%)
Maximum test specificity plus sensitivity threshold		
MPA only	567 (76.8%)	171 (23.2%)
IBA only	454 (61.5%)	284 (38.5%)
IBA and MPA	416 (56.4%)	322 (43.6%)

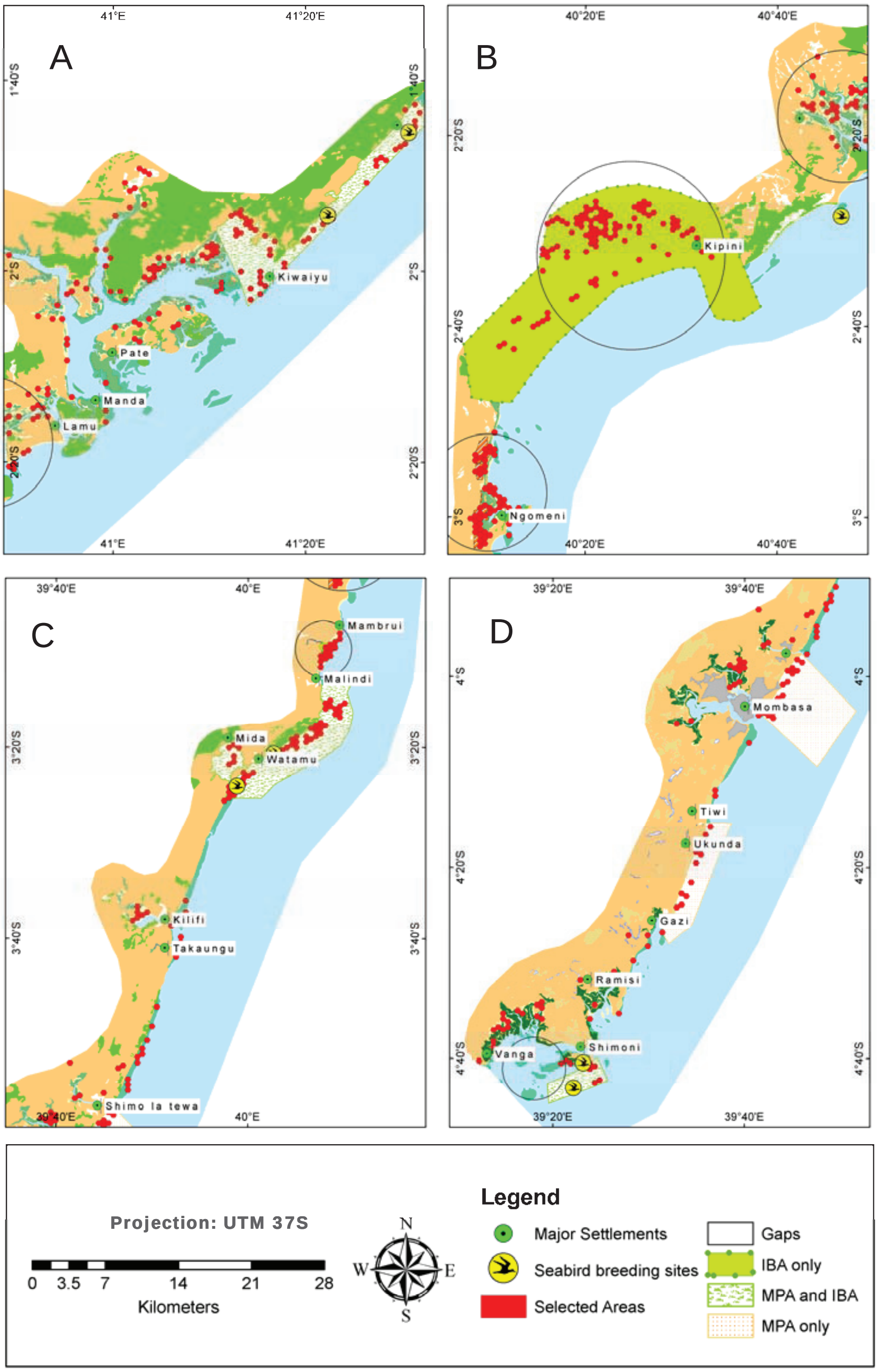


Figure 6: Marxan output when run with no restrictions (Equal sensitivity and specificity threshold)

3.4 Priority sites for conservation of seabirds and shorebirds that are excluded from the MPA network (Gaps)

When Marxan was run with MPAs locked in, clusters (priority areas) were identified around:

- Lamu Archipelago (west of Lamu - Figure 7: A,B)
- Tana River Delta (east of Kipini - Figure 7: B)
- Ngomeni area (east of Kipini - Figure 7: B)
- Sabaki River Mouth (south of Mambrui - Figure 7: C)
- Kisite area (south east of Shimoni - Figure 7: D)

When Marxan was run with MPAs and IBAs locked in, clusters (priority areas) were identified around:

- Lamu Archipelago (west of Lamu - Figure 8: A, B). This includes a small section to the extreme north of pate and north-west of Kiwaiyu.
- Ngomeni area (east of Kipini - Figure 8: B)

In summary, five high priority areas were not represented within the current protected areas (as identified when Marxan was run with MPAs locked in). Of these, three have already been recognized as high priority through the IBA process (Tana River Delta, Sabaki River Mouth and Kisite area). Two areas (Lamu archipelago and Ngomeni areas) have no formal recognition as being important for birds.

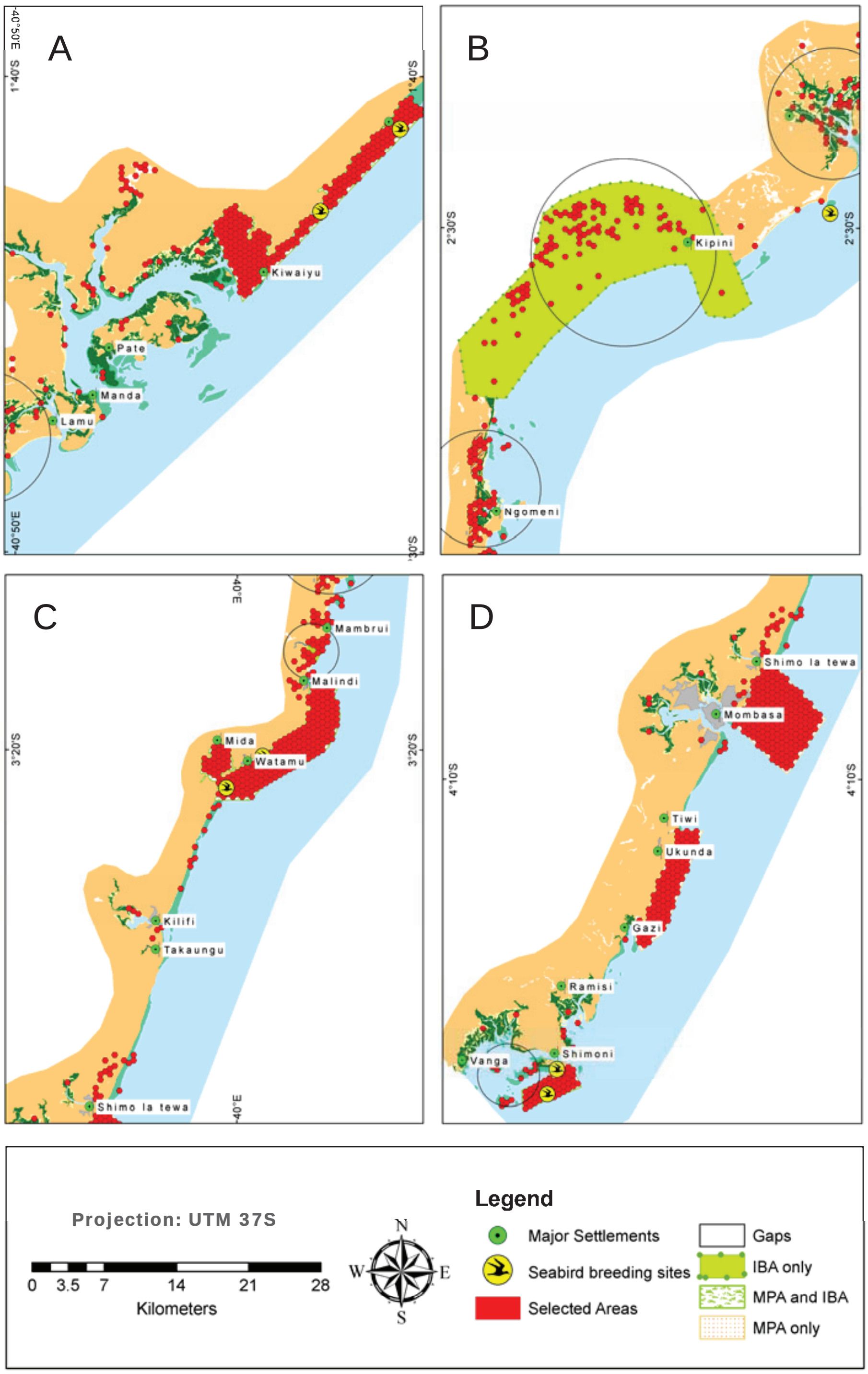


Figure 7: Marxan output when run with MPAs locked in

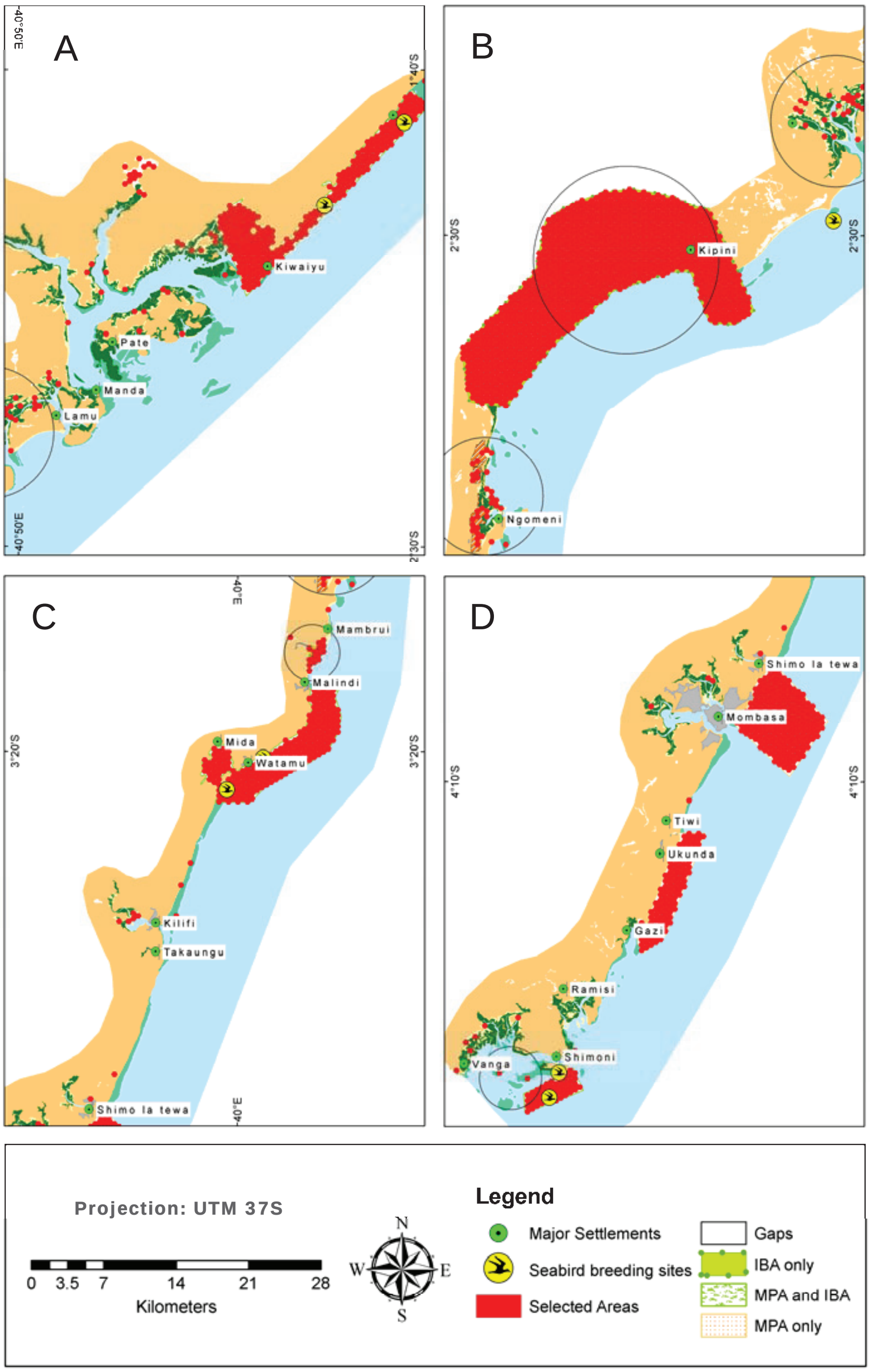


Figure 8: Marxan output with MPAs and IBAs locked in

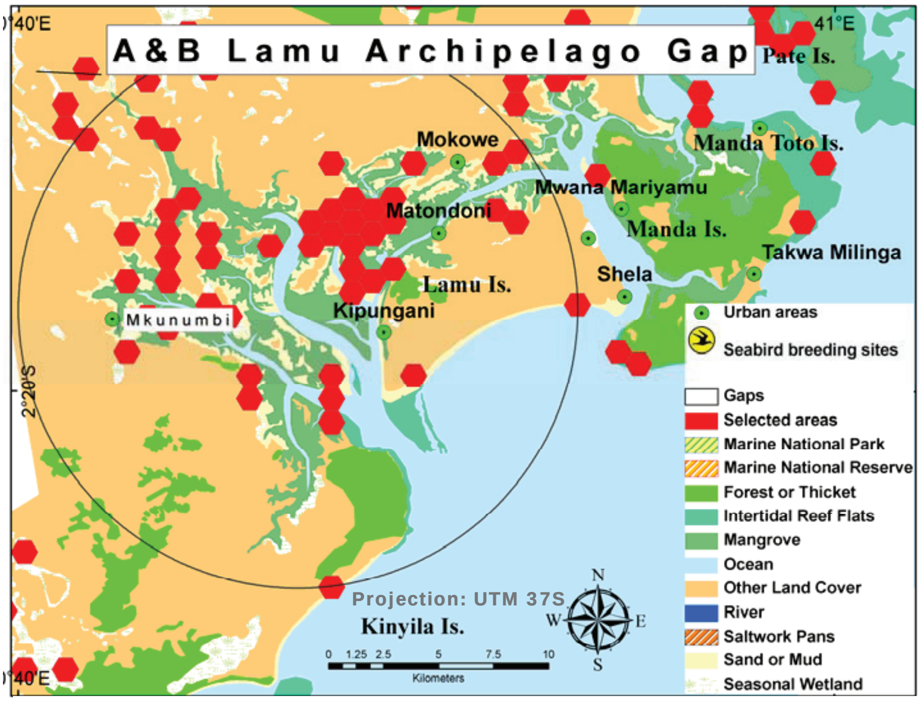


Figure 9: Identified gap in the Lamu Archipelago area

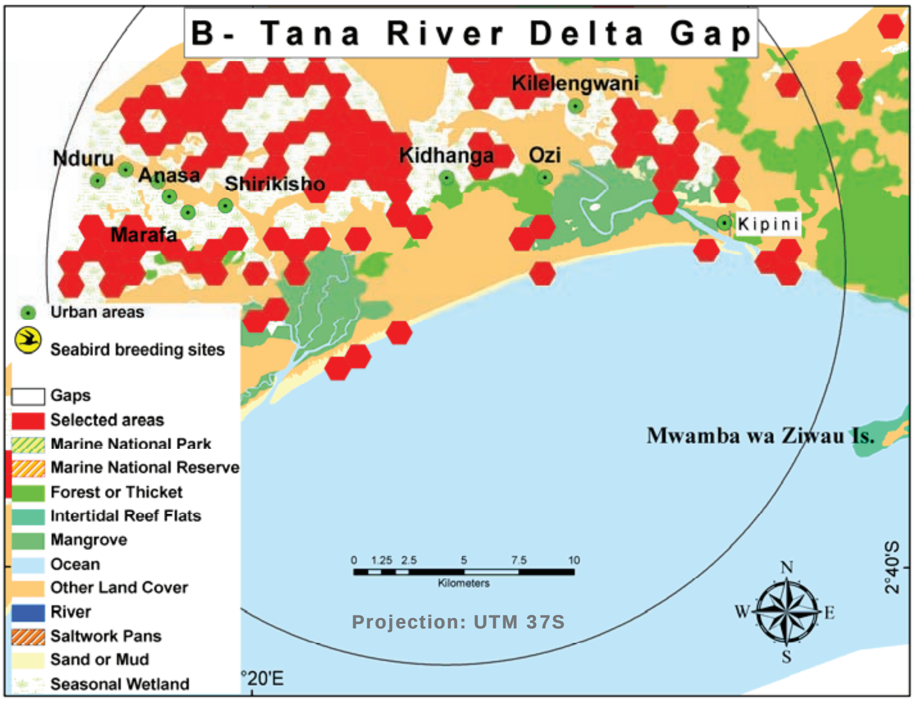


Figure 10: Identified gap in the Tana River delta area

The Lamu Archipelago (Figure 6-8: A, B and Figure 9) is composed of some areas located north of Kipungani and some located north of Mkunumbi. It is composed mainly of sand or mud flats within mangrove covered creeks (Figure 9).

The Tana River Delta (Figure 6-8: B and Figure 10) is composed mainly of seasonal wetlands located north-east of Kipini. The sandy beaches and the nearby island of Mwamba Ziwaiyu (Figure 10) are important roosting areas for seabirds and shorebirds.

The Ngomeni area (Figure 6-8: B and Figure 11) is composed of mangrove creeks and sand or mud flats. A significant area is also covered by salt extraction pans (Figure 11) run by different companies such as Kensalt.

The Sabaki River mouth area (Figure 6-8: C and Figure 11) is composed of muddy and sandy flats where the Sabaki river enters the ocean. The area is also covered with sand dunes.

The Kisite area (Figure 6-8: D and Figure 12) is composed of several areas of intertidal reef flats especially south of Wasini Island and around Sii island as well as the muddy/sandy flats north of Kaufumbani Island (Figure 12). While this area already has two MPAs (Kisite MNP and Mpunguti MNR), the sections identified fall outside the MPA boundaries.

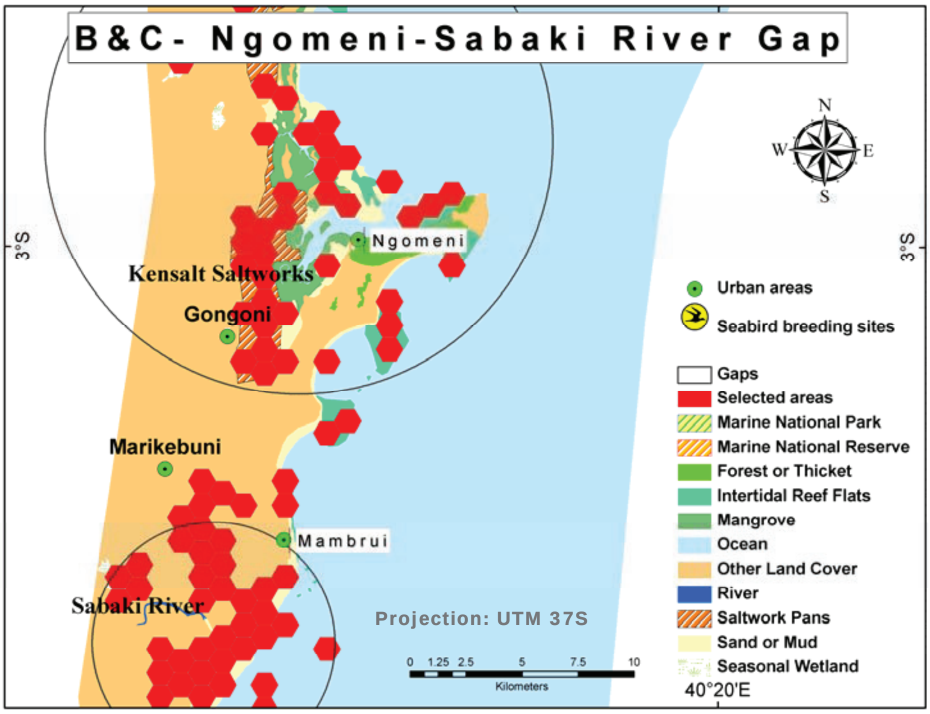


Figure 11: Identified gap in the Ngomeni and Sabaki River area

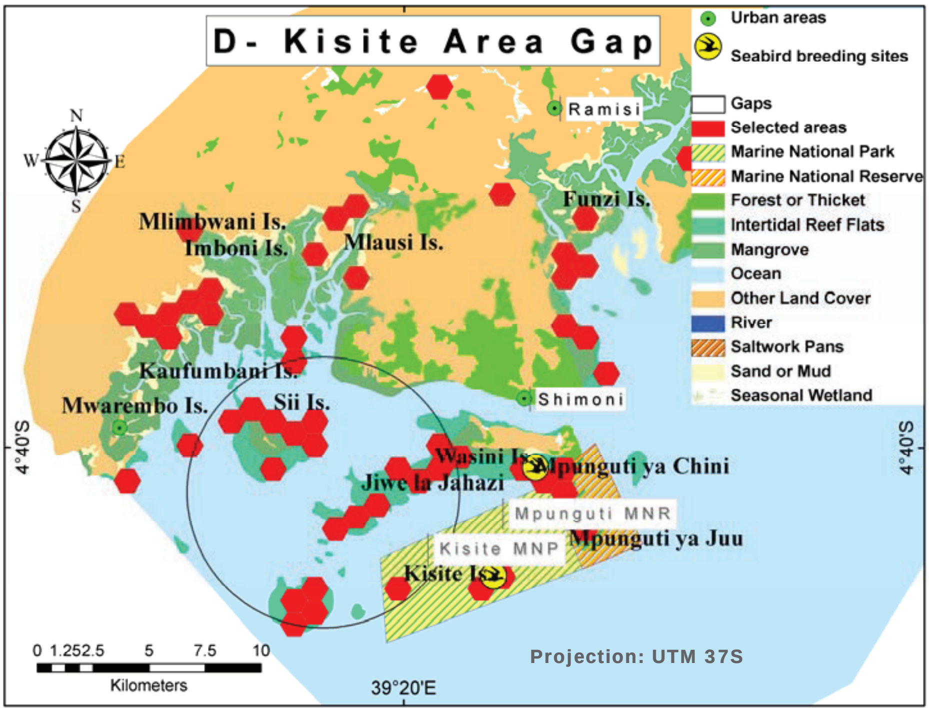


Figure 12: Identified gap in the Kisite area

4. Discussion

4.1 Maxent model validation

All the species modelled had very high AUC values, even those that were modelled with a few points. This can be attributed to the fact that AUC values tend to be higher for species that have narrow ranges compared to the area covered by the environmental data (Phillips, 2005). Most of the seabirds and shorebirds studied were habitat specific with most birds preferring less than four habitat types (Figure 3: A, C, D; Figure 5: A, B). Those that were even more specific in their habitat preference (occurring in 3 or less habitat types) had the highest AUC values. The high accuracies can also be attributed to the level of detail of the land cover layer. Since, this layer was manually digitized, the delineation of the various habitat types was more accurate thus minimizing the chances of a species presence point overlaying on the wrong habitat type. Nevertheless, in addition to validation using AUC values, a careful examination of the output maps was done with regard to comparison with known atlas distributions and likely habitat types. Previous studies that have combined the use of Maxent and Marxan have also had similarly high AUC values (Esselman & Allan, 2011; Urbina-Cardona & Flores-Villela, 2010)

4.2 Environmental and human factors influencing distribution of seabirds and shorebirds

Land cover was the most important factor in determining the distribution of seabirds and shorebirds in Kenya. This is expected given that these birds tend to be habitat specific. The implication is that an accurate and detailed cover layer is necessary to accurately predict the distribution of these species.

Most of the factors did not influence seabirds and shorebirds differently with the exception of distance to rivers. Distance to rivers had a stronger influence on seabirds and shorebirds. This could be due to the fact that deltaic waters such as that of the Tana River delta tend to be richer in fish resources on which the seabirds rely. Although the influence of other factors was not statistically significant, examining the response curves showed some differences in how seabirds and shorebirds react to factors. Several seabird species, for example, were found in seasonal wetlands while shorebirds were proportionally less. This could be due to the fact that the seabirds exploit the fish resources in these wetlands. There were also differences in response to distance from mangroves with most shorebirds having the highest probabilities of occurrence near the mangroves and some seabirds having a double peak. The second peak indicates probability of occurrence further away from the mangroves which is an indication of seabirds occupying remote islands which do not have mangrove stands. This is expected given the large foraging ranges of seabirds.

Seabirds and shorebirds had high probabilities of presence about 5 km from roads and towns. This is an indication how accessibility influences the ease of data collection and therefore its availability. There will be more data available for easily accessible areas than those that are remote and difficult to reach. Distance from MPAs generally showed a higher probability of species occurrence close to MPAs but some species had a second peak further away from the MPAs. While the first peak closer to the MPA could indicate the favourable effects of the protective nature of the MPAs, the second peak could be an indicator of a gap existing in the MPA system for those species.

Although studies were not done on what factors determine where seabirds will breed, observations during fieldwork showed that seabirds can shift breeding locations from more protected areas to less protected or unprotected areas. For example, in the 2011 breeding season, Roseate Terns that usually breed in the Whale Island (which is within a Marine National Park) shifted and attempted to breed in Darakisi off Watamu. This is a marine reserve with a lower degree of protection. As a result, most of the colony was destroyed by egg collectors (Colin Jackson pers. Comm.).

Figure 13: Photo of boys collecting tern eggs at Darakisi (Colin Jackson)

4.3 Priority sites for conservation of shorebirds and seabirds

Results from running Marxan without restrictions showed that the current IBAs and MPAs combined cover less than half of the high priority areas selected by Marxan thus indicating significant gaps in the protected area system. In fact, when MPAs were considered alone, only about a quarter of the high priority areas were represented. When Marxan was rerun with MPAs and MPAs plus IBAs locked in, the desired target could not be met using planning units within these areas in both cases. Therefore, both current MPAs and IBAs may not be sufficient to adequately represent shorebirds and seabirds in Kenya implying that additional areas need to be considered.

4.4 Priority sites for conservation of seabirds and shorebirds that are excluded from the MPA network (Gaps)

Lamu Archipelago and Ngomeni area were identified as high priority areas that were neither MPAs nor IBAs. The results show in order to meet the targeted area for conservation (i.e. 40% of distribution ranges and 100% of the breeding areas), these two areas need to be included within the reserve network. For the Lamu Archipelago area, there is no published record of a survey for seabirds and shorebirds available but there is some anecdotal evidence of seabirds breeding here (Zimmerman, *et al.*, 1999). Lack of credible data is probably responsible for its non-recognition as an important area for birds. It is therefore important to carry out surveys in this area to find out what it holds in terms of biodiversity. This is crucial especially because building of a second Kenyan port is soon to start here. This multi-billion shilling project will involve building of a deep-water harbour, high speed rails, oil pipelines, an international airport and an oil refinery (The East African, 2009).

There is also limited data for the Ngomeni area and especially on the salt pans. These are man-made ponds used for salt extraction where sea water is pumped into a series of ponds and water is allowed to evaporate leaving behind salt residues that are then collected. It is probable that the birds feed in the creeks and intertidal areas during low tide and then move to the saltworks during high tide to roost (Colin Jackson pers. comm). Whether they do any feeding at the saltworks requires further research. Saltworks present a conservation challenge, first because they are privately owned and second because large areas of mangroves are cleared during establishment of the ponds. So while seabirds and shorebirds might gain additional habitat, other biodiversity may be losing critical habitats. It is important to determine how the birds use the saltworks both spatially and temporally.

Other gaps identified (i.e. Sabaki River Mouth, Tana Delta and some parts of Kisite) are already designated as IBAs through a number of criteria developed by experts (Bennun & Njoroge, 1999). In the case of Kisite, part of the area is protected by the Kisite Marine National Park and the adjacent Mpunguti Marine Reserve (Figure 12) but a lot of the intertidal reef flats identified are outside these boundaries. Personal observation during fieldwork and anecdotal evidence from fishermen indicates that they could be important as roosting areas. Low tide surveys during the period with the highest number of seabirds/shorebirds (i.e. around January-Feburary) need to be carried out to investigate this.

Verifying the accuracy of Marxan's selection of high priority areas is difficult without additional fieldwork to these areas. However, the fact that there was an agreement between output from Marxan and areas identified by experts through the IBA process generates a degree of confidence in the results.

Of the areas identified as gaps, only the Tana River Delta and Sabaki River mouth are monitored frequently. Priority is therefore to ensure monitoring of the other areas as well. This is especially crucial for Lamu Archipelago and Ngomeni area since it is necessary to confirm whether these are high priority areas as determined by Marxan. This will also help to determine whether they meet the required thresholds to be designated as IBAs.

4.5 Challenges and limitations of the study

4.5.1 Data Availability and Quality

The nature of the study area made it difficult to get adequate data. Since seabirds and shorebirds use habitats located both on land and water, it was difficult to get datasets that cover both areas. For example, there were no Bioclim layers covering the intertidal areas and these had to be interpolated. For elevation, high resolution data was available for the land side and low resolution for the bathymetry and these had to be combined before they could be used in Maxent. There were also no land cover maps of the intertidal areas and these had to be created. For the species data, most of the secondary data that was used came from a survey of seabirds and shorebirds done in 1995 (Nasirwa*, et al.*, 1995) and which covered about half of the Kenyan coast. To my knowledge, this is the only systematic survey of the coast that has been done to date. Most monitoring activities are done on a few well known birding areas. Most areas are therefore unsurveyed, poorly surveyed or the data available is out-dated. The timing of the fieldwork for this study was also inappropriate given that most of the species modelled are migratory species and only arrive in Kenya in large numbers around January and February. Secondary data was therefore crucial in supplementing the field data. This lack of regular monitoring data raises a potential bias in terms of areas selected as high priority for conservation. This is because species presence or detection will depend on the season, time of day, tide level, sampling design, area searched and time spent searching among others. Therefore, if a species is recorded in a particular site, it is not possible to determine if this is its habitual site or whether it is there by coincidence unless regular monitoring is done.

4.5.2 Setting of thresholds and targets

A drawback of Marxan is that it does not use probability of occurrence as an input and therefore a threshold or cut-off point has to be determined to indicate a presence. If the wrong threshold is used, it may affect the planning units selected by Marxan. Results from this study show that when a lower threshold is used, the area of the identified priority sites increases considerably and this has implications on costs of establishing the reserves. Using the two threshold methods however identified similar priority areas and proportions in and out of MPAs. Therefore, thresholds should be chosen with care depending on desired objectives. It is therefore important to carry out a sensitivity analysis to determine the effect of varying thresholds on the final selected outcome. A sensitivity analysis is also required when setting a target of how much of the species distribution range you need to conserve. Ideally, setting targets needs input from experts who have studied the species (Ardron, *et al.*, 2010). Different species may require different targets based on their threat status, population trend or other issues peculiar to the species such as breeding patterns (Ardron, *et al.*, 2010). It is however recommended that at least 30% of a species' range should be protected (IUCN, 2003).

4.6 Implications of the study

The logical step after identifying gaps is to attempt to have them given some level of protection. Setting up of protected areas is becoming more difficult due to competing land uses and increased awareness about community rights. It is unlikely that any protected areas in future can be created based on only seabirds and shorebirds. It is therefore important to do a comprehensive systematic planning exercise that includes all biodiversity. This study demonstrates the importance of embracing the use of species distribution modelling (SDM), gap analysis and reserve design in the setting up and management of marine protected areas in Kenya. SDM can help solve issues of data gaps and issues of poor resolution data. Marxan is able to handle competing multiple objectives, for example, it is able to select the cheapest areas based on cost (e.g. cost of relocation, opportunity cost etc) but ensure all set targets are met (Ardron, *et al.*, 2010; Ball & Possingham, 2000). It is also possible to investigate the effect of negotiated settlements in terms of meeting the desired targets.

It is also important to reconsider the way in which MPAs are created and managed. For species such as seabirds and shorebirds which are highly mobile, traditional fixed reserves are not enough as has been demonstrated by seabirds varying their breeding locations to areas outside MPAs. There is scope for other protection options such as seasonal MPAs as well as breeding habitat improvement to ensure the birds breed within the protected areas.

5. Conclusions and Recommendations

5.1 Conclusions

This study found out that marine protected areas in Kenya do not adequately represent high priority areas for conservation of seabirds and shorebirds. Only about 20% of selected priority areas were found within MPAs. There were five high priority areas that were identified as gaps. Three of them (Tana River Delta, Sabaki River Mouth and Kisite area) have already been identified as Important Bird Areas (IBAs). Two of the areas (Lamu Archipelago and Ngomeni area) are not identified as IBAs and therefore field surveys are needed to determine whether they meet the required thresholds to be designated as IBAs. The major factor that influenced distribution of seabirds and shorebirds was cover. With the exception of distance from a river, all other factors did not significantly influence seabirds and shorebirds differently. This study demonstrates the usefulness of species distribution modelling and gap analysis techniques to conservation area planning.

5.2 Recommendations

There is need for field surveys to verify whether the areas selected by Marxan as high priority areas hold significant numbers of different species of seabirds and shorebirds as predicted. This can act as a validation method for Marxan results.

An investigation to find out which factors influence the choice of breeding locations by seabirds in Kenya is required. This is because any conservation measures need to protect not only the current breeding areas but also potential breeding areas.

To improve on modelling using Maxent, there is need for a more accurate map of the intertidal habitats at the Kenyan coast. In this study, the intertidal zone was treated as a single homogeneous habitat but in reality it is composed of different zones such as sandy intertidal, reef rock, coral gardens, seagrass beds among others.

There is also need for more data on seabirds and shorebirds. This requires setting up a monitoring programme for these species as well as developing a platform for sharing such data.

References

Ardron, J. A., Possingham, H. P., & Klein, C. J. (Eds.). (2010). *Marxan Good Practices Handbook, Version 2*. Victoria, B.C: Pacific Marine Analysis and Research Association.

Ball, I. R., & Possingham, H. P. (2000). Marxan (V1.8.2): Marine Reserve Design Using Spatially Explicit Annealing, a Manual.

Bennun, L., & Njoroge, P. (1999). *Important Bird Areas in Kenya*. Nairobi, Kenya: Nature Kenya, the East Africa Natural History Society.

Birdlife International. (2011a). Important Bird Areas factsheet: Kiunga Marine National Reserve. Retrieved 24/08/2011, from http://www.birdlife.org

Birdlife International. (2011b). Kenya's Tana River delta under seige. Retrieved 24/08/2011, from http://www.birdlife.org/news/news/2009/12/tana_update.html

Cantor, S. B., Sun, C. C., Tortolero-Luna, G., Richards-Kortum, R., & Follen, M. (1999). A Comparison of C/B Ratios from Studies Using Receiver Operating Characteristic Curve Analysis. *Journal of Clinical Epidemiology, 52*(9), 885-892.

Carnes, A., & Ogneva-Himmelberger, Y. (2011). Temporal Variations in the Distribution of West Nile Virus Within the United States; 2000–2008. *Applied Spatial Analysis and Policy*, 1-19.

Delavenne, J., Metcalfe, K., Smith, R. J., Vaz, S., Martin, C. S., Dupuis, L., Coppin, F., & Carpentier, A. (2011). Systematic conservation planning in the eastern English Channel: comparing the Marxan and Zonation decision-support tools. *ICES Journal of Marine Science: Journal du Conseil, 69*(1), 75-83.

Esselman, P. C., & Allan, J. D. (2011). Application of species distribution models and conservation planning software to the design of a reserve network for the riverine fishes of northeastern Mesoamerica. *Freshwater Biology, 56*(1), 71-88.

Francis, J., Nilsson, A., & Waruinge, D. (2002). Marine Protected Areas in the Eastern African Region: How Successful Are They? *AMBIO: A Journal of the Human Environment, 31*(7), 503-511.

Government of Kenya. (2009). *State of the coast report: towards integrated management of coastal and marine resources in Kenya*. Nairobi: National Environment Management Authority (NEMA).

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling, 135*(2-3), 147-186.

IUCN. (2003). *Recommendations of the Fifth IUCN World Parks Congress, Durban. Benefits Beyond Boundaries*. Retrieved from http://cmsdata.iucn.org/downloads/recommendationen.pdf.

Johnson, C. J., & Gillingham, M. P. (2005). An evaluation of mapped species distribution models used for conservation planning. *Environmental Conservation, 32*(02), 117-128.

Lawler, J. J., Wiersma, Y. F., & Huettmann, F. (2010). Using species distribution models for conservation planning and ecological forecasting. In C. A. Drew, Y. F. Wiersma & F. Huettmann (Eds.), *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*. New York: Springer Verlag.

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography, 28*(3), 385-393.

Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature, 405*(6783), 243.

Marini, M., Barbet-Massin, M., Lopes, L., & Jiguet, F. (2010). Predicting the occurrence of rare Brazilian birds with species distribution models. *Journal of Ornithology, 151*(4), 857-866.

Nasirwa, O., Oyugi, J., Jackson, C., Lens, L., Bennun, L., & Seys, J. (1995). *Surveys of waterbirds in Kenya, 1995: Lake Victoria wetlands, south Kenya coast and Tana River dams.* Nairobi: National Museums of Kenya.

O'brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity, 41*(5), 673-690.

Phillips, S. J. (2005). A brief tutorial on Maxent. from http://www.cs.princeton.edu/~schapire/maxent/turorial/tutorial.doc.

Rodríguez, J. P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions, 13*(3), 243-251.

Rondinini, C., Wilson, K. A., Boitani, L., Grantham, H., & Possingham, H. P. (2006). Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters, 9*(10), 1136-1145.

Scott, J. M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., Derchia, F., Edwards, T. C., Ulliman, J., & Wright, R. G. (1993). Gap analysis- A geographic approach to protection of biological diversity. *Wildlife Monographs, 123*, 1-41.

Sokal, R. R., & Oden, N. L. (1978). Spatial autocorrelation in biology: 1. Methodology. *Biological Journal of the Linnean Society, 10*(2), 199-228.
Stewart, R. R., Noyce, T., & Possingham, H. P. (2003). Opportunity cost of ad hoc marine reserve design decisions: an example from South Australia. *Marine Ecology-Progress Series, 253*, 25-38.
The East African. (2009). Lamu port set to be a growth catalyst. Retrieved 24/8/2011, from http://www.theeastafrican.co.ke/business/-/2560/610676/-/5k1t16z/-/index.html%5D
Tychsen, J. (Ed.). (2006). *KenSea. Environmental Sensitivity Atlas for Coastal Area of Kenya*. Copenhagen: Geological Survey of Denmark and Greenland (GEUS).
UNEP. (1998). *Eastern Africa atlas of coastal resources 1: Kenya*. Nairobi: United Nations Environment Programme (UNEP).
Urbina-Cardona, J. N., & Flores-Villela, O. (2010). Ecological-Niche Modeling and Prioritization of Conservation-Area Networks for Mexican Herpetofauna. *Conservation Biology, 24*(4), 1031-1041.
Watts, M. E., Ball, I. R., Stewart, R. S., Klein, C. J., Wilson, K., Steinback, C., Lourival, R., Kircher, L., & Possingham, H. P. (2009). Marxan with Zones: Software for optimal conservation based land and sea use zoning. *Environmental Modelling & Software, 24*(12), 1513-1521.
Wells, S., Burgess, N., & Ngusaru, A. (2007). Towards the 2012 marine protected area targets in Eastern Africa. *Ocean and Coastal Management, 50*(1-2), 67-83.
Zimmerman, D. A., Turner, D. A., & Pearson, D. J. (1999). *Birds of Kenya and Northern Tanzania*. London: Christopher Helm.

Appendices

Appendix 1: Data and software sources

Images	**Resolution**	**Website**
Landsat 5 TM	30m	http://glovis.usgs.gov/
Aster GDEM	30m	http://www.gdem.aster.ersdac.or.jp/
ETOPO 1	1.82 km	http://ngdc.noaa.gov/mgg/global/global.html
Bioclim	1 km	http://www.worldclim.org/
Other data	**Type**	**Website**
Kensea Maps	PDF	http://mirror.undp.org/kenya/KenSea.htm
Base data	Shapefiles	http://www.wri.org/publication/content/9291
Bird data	csv	http://www.worldbirds.org/v3/kenya.php
Tide Tables	Online	http://tides.mobilegeographics.com/locations/3076.html
Software	**Price**	**Website**
Marxan	Free	http://www.uq.edu.au/marxan/
Zonae Cogito	Free	http://www.uq.edu.au/marxan/
PAT Tools	Free	http://gg.usm.edu/pat/index.htm

Appendix 2: Land cover classes

Land Cover (Habitat Type)	**Code**
Ocean	1
Other land cover	2
River	3
Intertidal Reef Flats	4
Forest or Thicket	5
Seasonal Wetland	6
Sand or Mud	7
Urban areas, Towns and Settlements	8
Saltworks Pans	9
Mangrove	10

Appendix 3: Results of Multicollinearity and spatial autocorrelation analysis

	Collinearity	Spatial autocorrelation		
Factor	**VIF value**	**Morans I**	**Z-score**	**Type**
Isothermality	12.1	0.63	9.13	C
Mean Temp. Driest Qtr.	6.8	0.62	8.92	C
Precip. Wettest Month	4.3	0.23	3.66	C
Precipitation Seasonality	3.4	0.53	7.7	C
Precip. of Warmest Qtr.	2.5	0.31	4.69	C
Elevation	1.6	0.03	0.81	R
NDVI	1.6	-0.07	-0.55	R
Distance fish landings	2.6	0.13	2.74	C
Distance Hotels	10.1	0.29	4.7	C
Distance Mangroves	5.3	0.28	4.5	C
Distance MPAs	8.1	0.4	6.1	C
Distance rivers	5.1	0.17	2.97	C
Distance roads	2.4	0.25	4.21	C
Distance urban areas	2.9	0.06	1.38	R

C-Clustered, R-Random

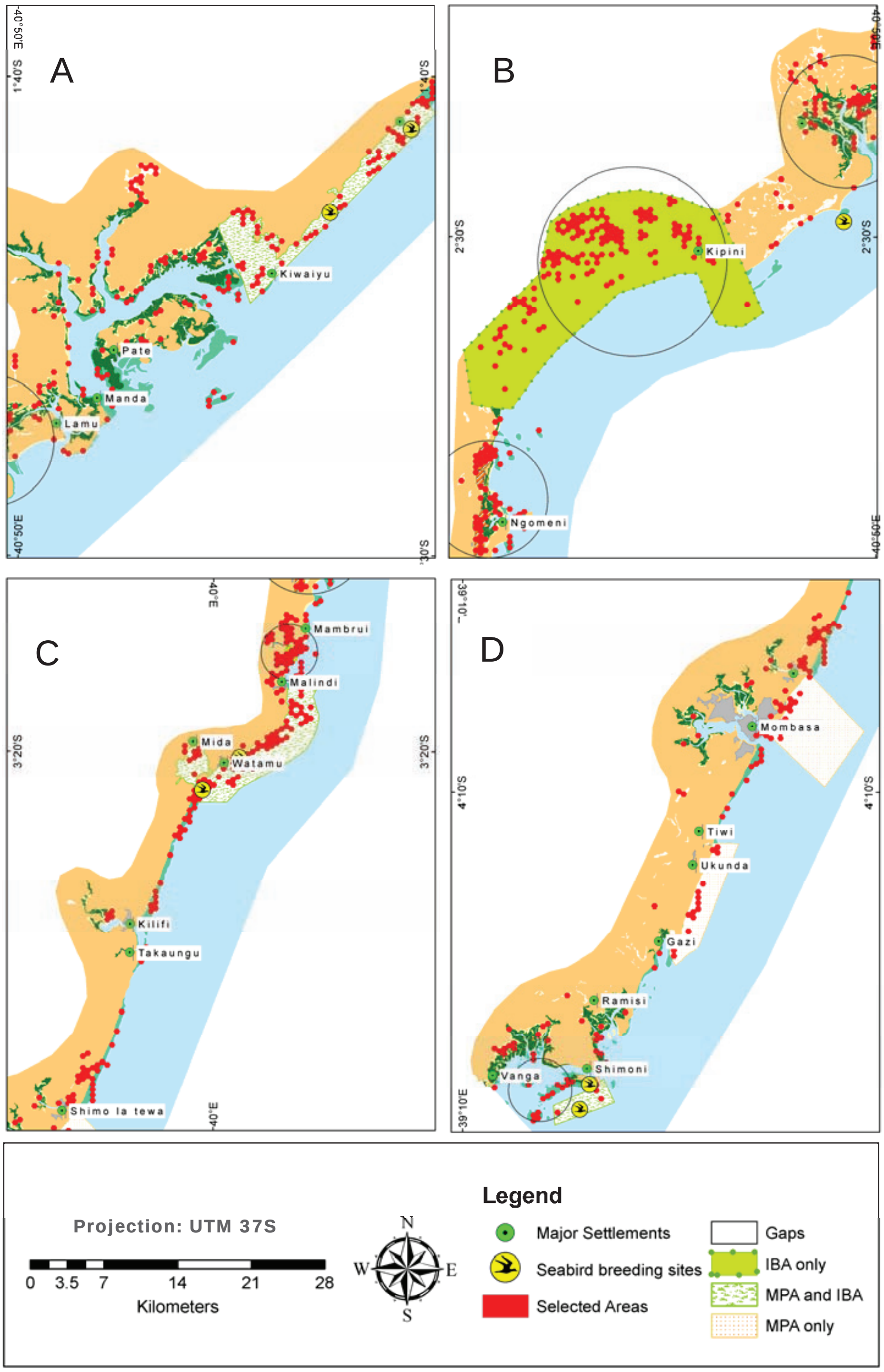
Appendix 4: Training gain values for seabirds with one factor used

Training gain with only variable	Caspian Tern	Common Tern	Greater Crested Tern	Gull-billed Tern	Lesser Crested Tern	Saunders's Tern	Sooty Gull
bio13	0.3	0.4	0.4	0.2	0.2	0.5	0.2
bio15	0.4	0.6	0.8	0.2	0.3	0.8	0.3
bio18	1.1	0.5	1.0	0.2	0.4	0.9	0.4
bio3	1.0	0.0	1.0	0.1	0.3	0.2	0.3
bio9	0.3	0.3	0.6	0.1	0.1	0.6	0.2
cover	1.6	3.0	3.1	2.3	2.2	2.9	2.8
dist_fish	1.1	1.3	0.5	0.3	0.7	0.8	1.1
dist_hotel	1.9	2.5	1.2	0.7	1.1	1.7	1.5
dist_mangr	0.1	0.4	0.4	0.2	0.1	0.4	0.1
dist_mpa	0.6	1.3	0.6	0.5	0.3	0.8	0.6
dist_river	1.0	0.6	1.8	0.4	0.9	0.9	0.7
dist_road	1.2	1.2	1.2	0.3	0.6	1.4	0.8
dist_urban	1.1	1.4	0.9	0.8	0.7	1.2	0.8
elevation	0.9	1.2	0.9	0.7	1.0	0.9	1.1
ndvi-raw	0.3	0.7	0.7	0.5	0.5	0.8	0.6

Appendix 5: Training gain values for shorebirds with one factor used

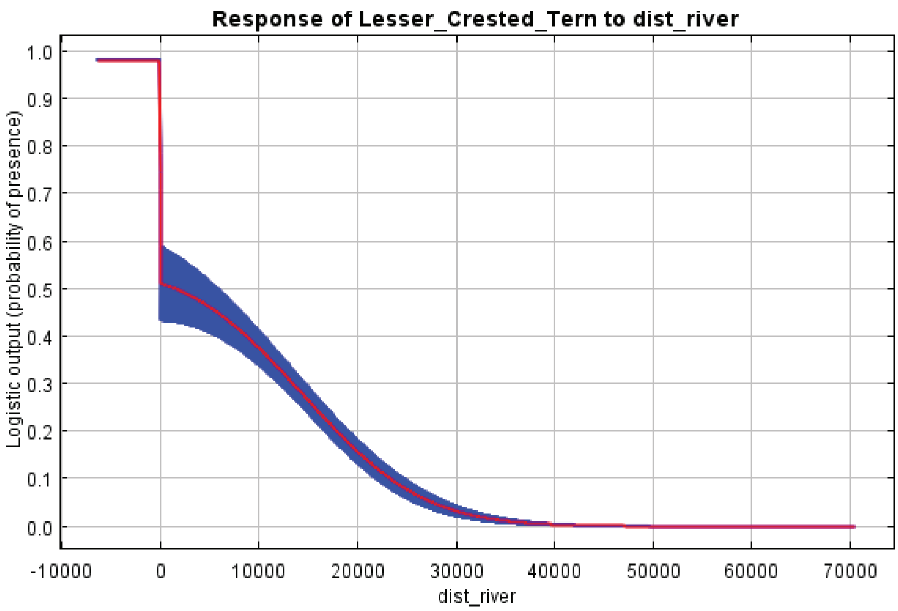
Training gain with only variable	**Grey Plover**	**Lesser Sandplover**	**Ruddy Turnstone**	**Sanderling**	**Terek Sandpiper**	**Whimbrel**	**White-fronted Plover**	**Crab Plover**	**Eurasian Curlew**	**Great Egret**	**Greater Sandplover**
bio13	0.1	0.4	0.4	0.3	0.1	0.0	0.3	0.1	0.3	0.0	0.1
bio15	0.1	0.5	0.3	0.3	0.1	0.1	0.7	0.1	0.4	0.1	0.1
bio18	0.2	0.8	0.7	0.5	0.0	0.1	0.7	0.1	0.4	0.2	0.0
bio3	0.1	0.3	0.3	0.1	0.1	0.0	1.1	0.1	0.2	0.3	0.0
bio9	0.1	0.5	0.1	0.1	0.0	0.1	0.6	0.1	0.4	0.1	0.0
cover	2.6	2.9	2.8	2.8	3.0	2.7	3.4	3.0	3.3	2.4	2.7
dist_fish	1.0	0.6	1.6	1.0	0.7	0.7	0.7	0.9	0.9	1.0	0.9
dist_hotel	1.1	1.7	2.5	2.1	0.5	0.8	0.8	0.5	1.1	1.1	1.0
dist_mangr	0.4	0.1	0.1	0.1	1.0	0.6	0.3	0.4	1.8	0.5	0.4
dist_mpa	0.5	1.1	1.0	1.2	0.3	0.5	0.3	1.0	0.5	0.6	0.5
dist_river	0.2	0.8	0.2	0.1	0.4	0.2	1.6	0.1	0.7	0.2	0.1
dist_road	0.6	1.2	1.1	0.9	0.3	0.4	0.9	0.4	1.0	0.4	0.5
dist_urban	0.8	0.9	0.9	1.1	0.3	0.7	0.7	0.9	1.1	0.5	1.0
elevation	0.8	1.1	1.1	1.2	0.7	0.8	0.8	0.7	0.8	0.8	0.9
ndvi-raw	0.4	1.1	1.0	1.0	0.3	0.5	0.3	0.3	0.8	0.3	0.4

Appendix 6: Marxan output when run with no restrictions (Maximum test sensitivity plus specificity threshold)



Appendix 7: Comparison of effect of distance from rivers on seabirds and shorebirds

A) Seabird example: Lesser Crested Tern



B) Shorebird example: Greater Sandplover