



# Assessing the relationship between environmental factors and malaria vector breeding sites in Swaziland using multi-scale remotely sensed data

Sabelo Nick Dlamini,<sup>1,2</sup> Jonas Franke,<sup>3</sup> Penelope Vounatsou<sup>1,2</sup>

<sup>1</sup>Swiss Tropical and Public Health Institute, Basel; <sup>2</sup>University of Basel, Switzerland;

<sup>3</sup>Remote Sensing Solutions GmbH, Baierbrunn, Germany

Correspondence: Penelope Vounatsou, Swiss Tropical and Public Health Institute, Socinstrasse 57, 4002 Basel, Switzerland.  
Tel. +41.61.2848109 - Fax: +41.61.2848101.  
E-mail: penelope.vounatsou@unibas.ch

Key words: Vector breeding; Remote sensing; Larva scooping; Malaria; Swaziland.

Acknowledgements: the authors would like to thank the European Commission under the Seventh Framework Programme (FP7) for partially funding the project. Some EO data were provided by the ESA managed GSC-DA, funded under ESA-EC Agreement on the Implementation of the Space Component of Global Monitoring for Environment and Security (GMES). We thank the National Malaria Control Programme (NMCP) of Swaziland for conducting and coordinating the field data collection, the NMCP's senior entomologist Mr. Quinton Dlamini for leading the field team during the entomological survey and NMCP manager Mr. Simon Kunene for approval to conduct the study using NMCP equipments. We would also like to acknowledge NMCP officer Mr. Sibonakaliso Vilakati for technical assistance during the preliminary planning phase.

Contributions: SND, entomological data collecting, analyzing and manuscript writing; JF, SND, remote sensing data processing, analysis and manuscript reviewing; PV, SND, JF, statistical analysis and manuscript reviewing.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: the work was supported by the European Commission under the Seventh Framework Programme (FP7).

Note: the present paper includes results from MALAREO (2011-2013; [http://cordis.europa.eu/project/rcn/97892\\_en.html](http://cordis.europa.eu/project/rcn/97892_en.html)), a project in the Seventh Framework Programme of the EU, aiming to stimulate and facilitate the use of earth observation in malaria control and management in South-Africa, Swaziland and Mozambique.

Received for publication: 7 January 2015.

Revision received: 24 February 2015.

Accepted for publication: 3 March 2015

©Copyright S.N. Dlamini et al., 2015  
Licensee PAGEPress, Italy  
Geospatial Health 2015; 10:302  
doi:10.4081/gh.2015.302

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 3.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Abstract

Many entomological studies have analyzed remotely sensed data to assess the relationship between malaria vector distribution and the associated environmental factors. However, the high cost of remotely sensed products with high spatial resolution has often resulted in analyses being conducted at coarse scales using open-source, archived remotely sensed data. In the present study, spatial prediction of potential breeding sites based on multi-scale remotely sensed information in conjunction with entomological data with special reference to presence or absence of larvae was realized. Selected water bodies were tested for mosquito larvae using the larva scooping method, and the results were compared with data on land cover, rainfall, land surface temperature (LST) and altitude presented with high spatial resolution. To assess which environmental factors best predict larval presence or absence, Decision Tree methodology and logistic regression techniques were applied. Both approaches showed that some environmental predictors can reliably distinguish between the two alternatives (existence and non-existence of larvae). For example, the results suggest that larvae are mainly present in very small water pools related to human activities, such as subsistence farming that were also found to be the major determinant for vector breeding. Rainfall, LST and altitude, on the other hand, were less useful as a basis for mapping the distribution of breeding sites. In conclusion, we found that models linking presence of larvae with high-resolution land use have good predictive ability of identifying potential breeding sites.

## Introduction

Malaria is caused by *Plasmodium* parasites, which are transmitted to people through a bite by an infected female *Anopheles* mosquito. Most mosquito species oviposit in standing waters and pools of varied amounts and sizes depending on the preference of each particular species. Targeting mosquito larvae and pupae with larvicides in standing water or breeding sites is one of the most important intervention measures in the fight against, and elimination of, malaria (Clennon *et al.*, 2010; Dambach *et al.*, 2014). Identification and mapping of all potential vector breeding sites is a prerequisite for successful vector control, especially larval source management (LSM) applied for effective elimination of residual foci. A national inventory of all residual foci is necessary if transmission is to be interrupted and remaining foci cleared (Chanda *et al.*, 2013). Past statistical modelling and mapping efforts have predicted vector distributions at continental scales based on climatic suitability and low-resolution remotely sensed (RS) data (Kulkarni *et al.*, 2010). However, very few studies have used entomological data in conjunction with remotely sensed data to identify,

map and predict potential malaria vector breeding sites at explicit geographical locations (Ahmad *et al.*, 2011; Bøgh *et al.*, 2007; Li *et al.*, 2008). The spatial resolution of the satellite-generated imagery is crucial for identifying potential vector habitats, and high spatial resolutions must be applied in order to capture not only larger water bodies, but also smaller ones, which are potentially as important for breeding.

Previous studies using satellite-generated imagery to identify suitable vector habitats, also based their approach on existing knowledge on how factors, such as temperature, humidity and rainfall, influence mosquito population dynamics and distribution (Beck *et al.*, 2000). Dambach *et al.* (2009) used imagery from the SPOT-5 satellite with supervised classification to identify land cover types known to be suitable as *Anopheles* mosquito breeding sites. Since no field-generated data were used in the analysis, the classification of relative risk was entirely based on the literature on *Anopheles* mosquito presence in different land cover types. Oesterholt *et al.* (2006) approximated vector breeding sites by assessing malaria incidence in relation to the distance to the nearest water body using geographical information systems (GIS). In their study, 10 houses were mapped and light traps hung at the end of an occupied bed to catch mosquitoes. Traps were emptied and mosquito species counted and determined the following morning. The identification of potential breeding sites using adult mosquito entomological data relies on the ability and precision to map distance to the nearest water bodies rather than distance to the actual breeding sites, which can be assessed, for instance, through larval scooping. However adult mosquito dispersal could be influenced by wind speed and wind direction (Bøgh *et al.*, 2007) making it a challenge to approximate the breeding sites.

Strong progress in the fight against malaria has been made in Swaziland. In 2002, insecticide-treated bed nets were introduced to complement the ongoing indoor residual spraying (IRS) activities. In 2008, after 15 years of progressive reduction of the disease burden (from 4005 to 369 cases, the country was nominated to spearhead the malaria elimination in the Southern African Development Community, which is being pursued according to the strategic plan for the period 2008-2015; Malaria Indicator Survey, 2010). Although the country has

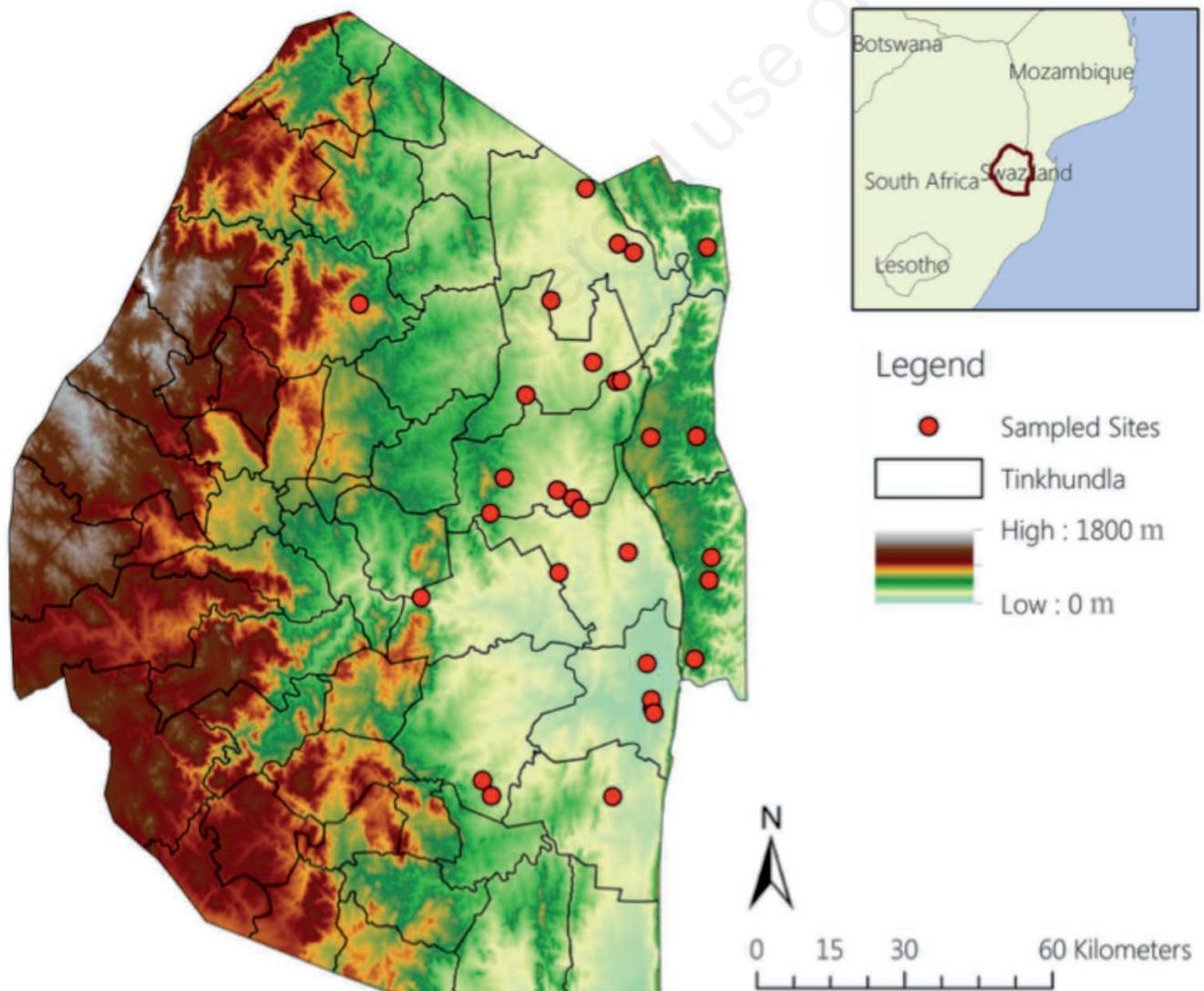


Figure 1. Altitude map showing the location of larva sampling sites.

consistently and annually applied IRS using DDT as its mainstay vector control intervention strategy, studies designed to support these attempts with empirical evidence on its effect on the number of vector breeding sites and their distribution have not been conducted. Identification and elimination of residual foci along with efforts to reduce the number of local malaria cases to zero remains a challenge, especially if there is a lack of geographically explicit supporting maps to target intervention efforts. Following the Stockholm Convention on Persistent Organic Pollutants ([http://sites.duke.edu/malaria/the-stock-](http://sites.duke.edu/malaria/the-stockholm-convention/)

[holm-convention/](http://sites.duke.edu/malaria/the-stockholm-convention/)), many countries will soon have very limited supply of DDT, so optimal ways to use this chemical in high priority areas must be sought. This is possible with spatially explicit maps guiding ground IRS activities, thus avoiding the indiscriminate use of DDT that leads to unnecessary waste and environmental damage in addition to potentially increased vector resistance.

The objective of this study was to analyze the relationship between environmental factors and malaria vector breeding sites in Swaziland by linking entomological data with multi-scale RS data and scooping for

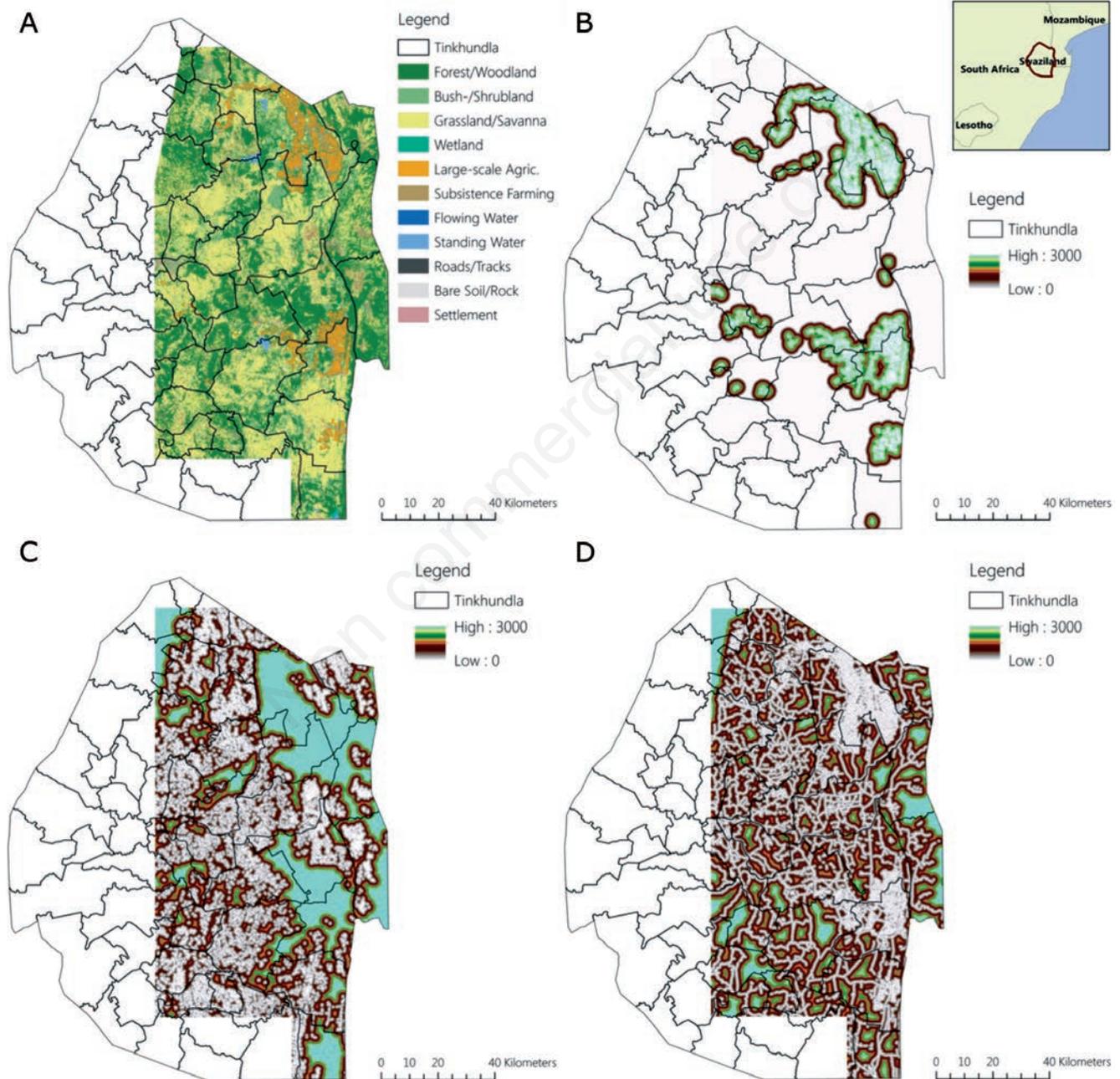


Figure 2. A) Land cover map covering the malaria-endemic area of Swaziland; B) distance-to-large-scale agriculture (m); C) distance-to-subsistence farming (m); D) distance-to-roads/tracks (m).

larvae in selected water bodies and dams, feeding collected information into a statistical regression model and using data mining tools to investigate potential associations. By this approach we aimed to contribute to the existing knowledge about malaria vector breeding habitats in Swaziland and provide high-resolution, spatially explicit maps to assist ongoing conventional control efforts as the country targets elimination by 2015.

## Materials and Methods

### Study area

Swaziland covers an area of 17,363 km<sup>2</sup> and consists of a mountainous Highveld (the wet western part of the country), which has an altitude of about 1800 m above the sea, and the relatively flat, dry eastern

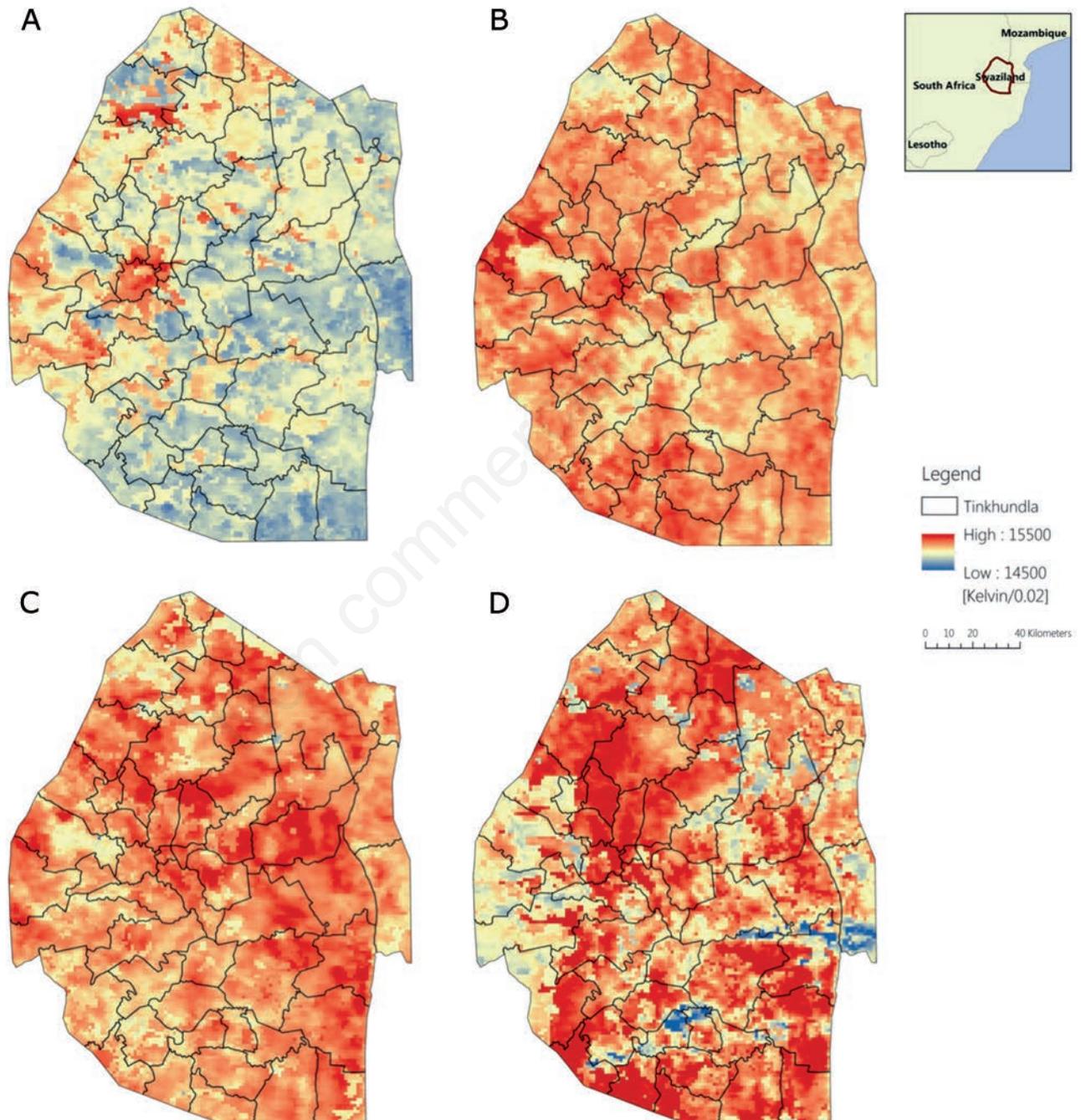


Figure 3. Land surface temperature: A) first week; B) second week; C) third week; D) fourth week.



part (between 100-300 m in altitude) called the Lowveld (Figure 1). The rain that falls on the Highveld flows towards the Lowveld, where it stagnates creating pools of standing water suitable for mosquito breeding because of the flat terrain and high temperatures in this dry zone.

Malaria transmission in Swaziland is prevalent along the eastern part of the country borders with Mozambique and north-east South Africa. Transmission occurs in the rainy season between November and May with a peak in February and March, sometimes extending to April. Transmission is unstable and follows the quantity of rainfall in

each particular year and the amount of standing water accumulated during the latest rain episode. The entire population is at risk as it generally lacks acquired immunity and is therefore highly vulnerable. The situation is particularly serious with respect to pregnant women and children under the age of 5 years (Malaria Indicator Survey, 2010). *Plasmodium falciparum* dominates and about 99% of malaria cases are infected with this species, while infections due to the other malaria species are occasional.

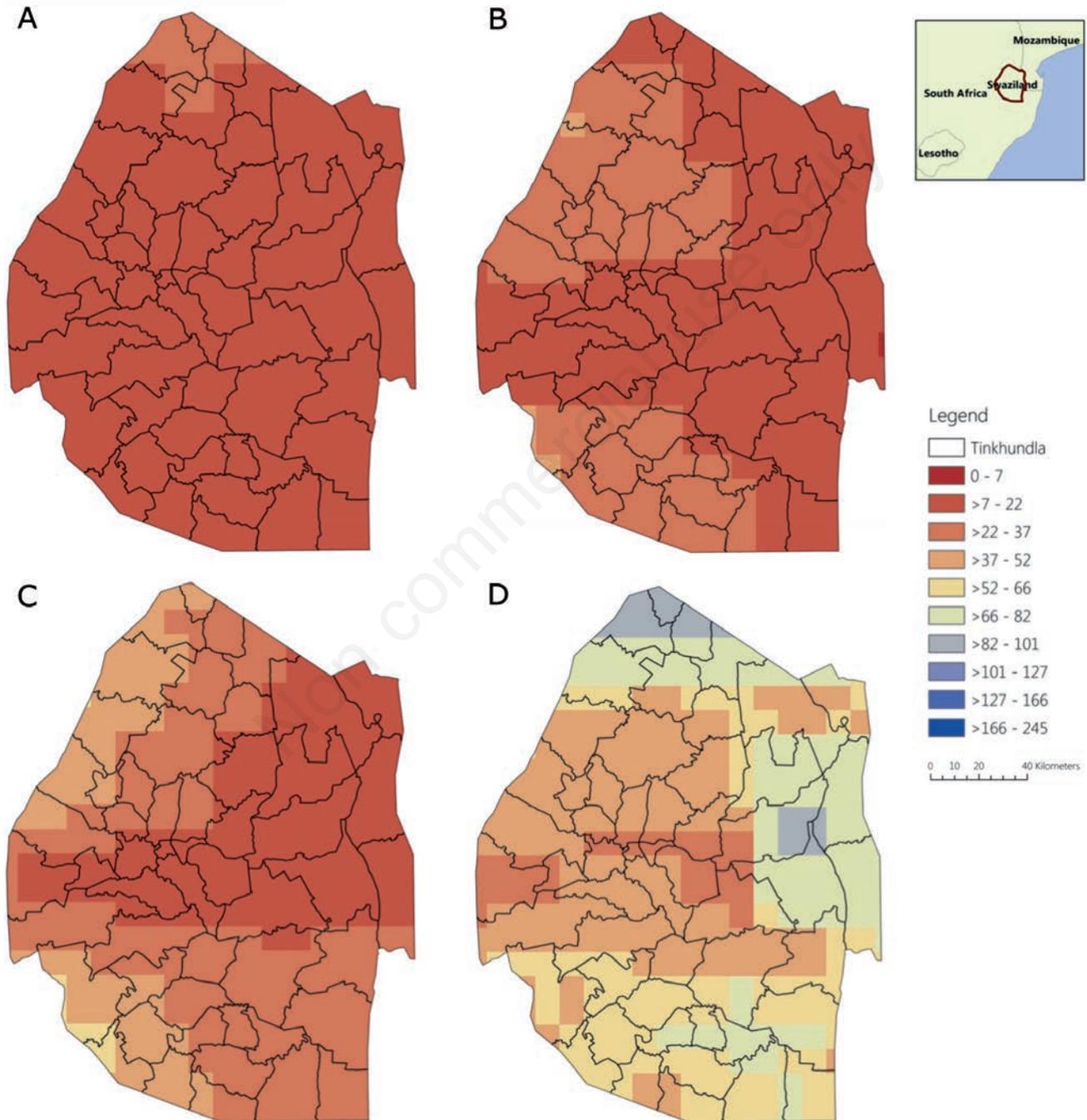


Figure 4. Data rainfall estimate (mm) of A) first 10-day period; B) second 10-day period; C) third 10-day period; D) fourth 10-day period.

## Entomological data

The National Malaria Control Programme (NMCP) of Swaziland recently developed a geo-database of all potential vector breeding sites and ongoing larva scooping activities. Data on vector control and entomology were extracted from the NMCP database and used in this study for an analysis including multi-scale RS data.

Mosquito larvae were collected from 30 selected water body/wetland sites and peripheral shallow pools of standing water as well as water-filled cattle hoof prints that were identified as potential breeding sites

in the Lowveld region. Figure 1 shows the larvae sampling locations in the different constituencies, locally known as Tinkhundla. The morphology of all visible larvae in the water was closely observed in an attempt to collect those larvae known to be anopheline. The larvae were scooped up from high-density larval areas in the water and the number of larvae per scoop counted. Where the larval density was extremely low, or no larvae visible, about 100 scoops were done and the density determined on the basis of the larvae present in each one. Following speciation by a senior entomologist, the mosquitoes belonging to the *An. gambiae* complex and *An. funestus* were kept in Eppendorf tubes,

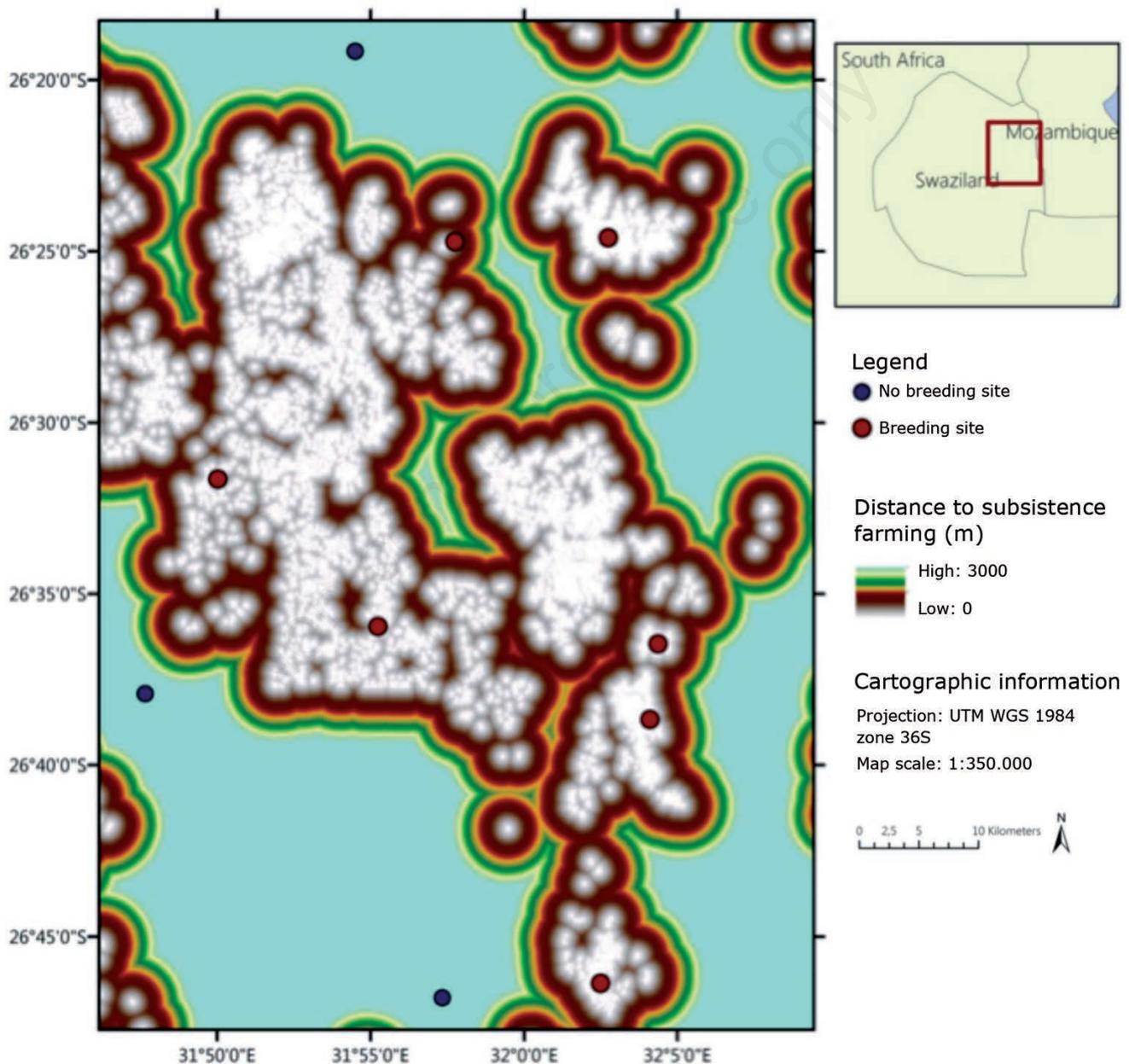


Figure 5. Example of the spatial distribution of some sampled vector breeding sites and the distance-to-subsistence farming in Swaziland.

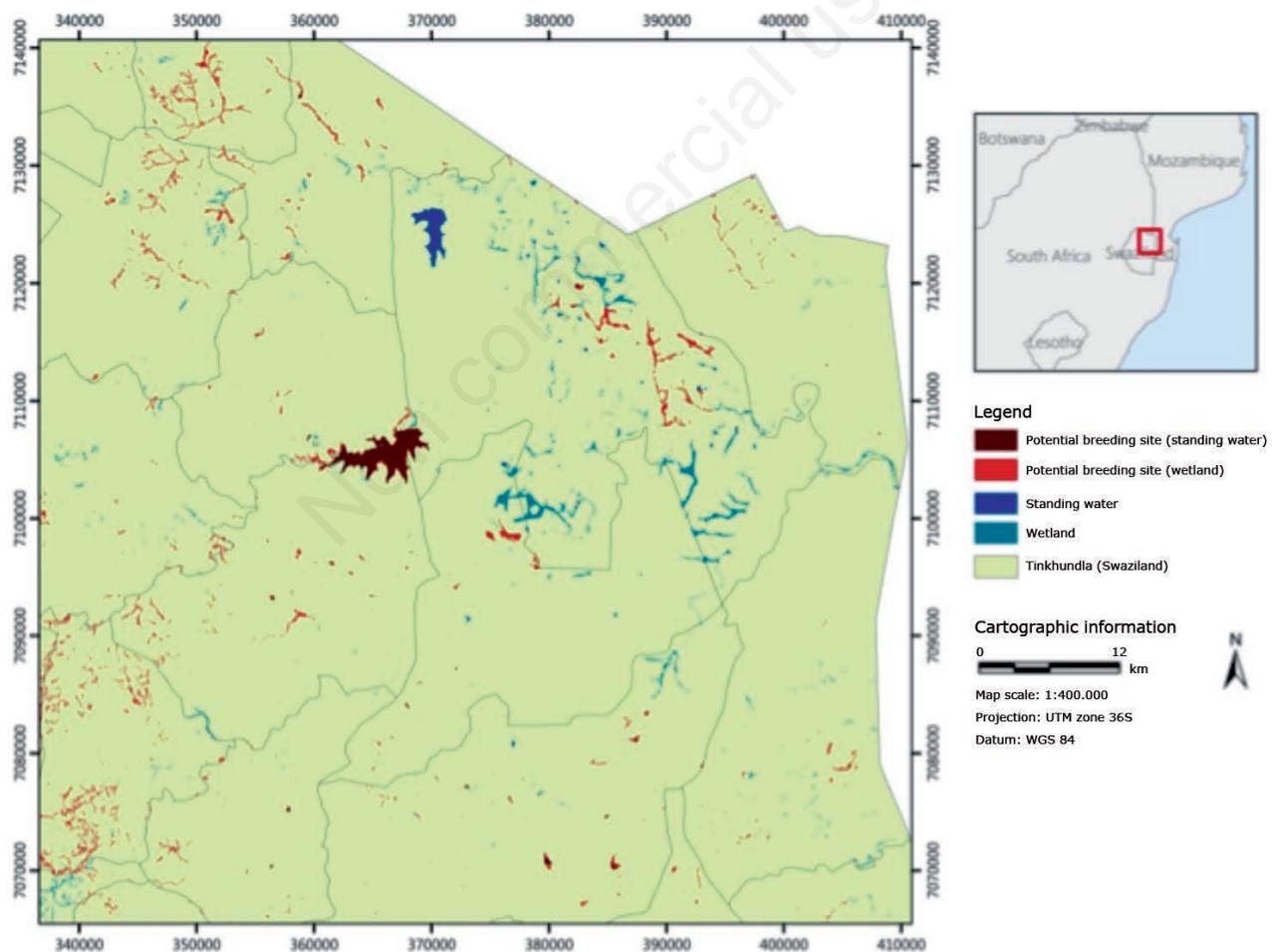
labeled and preserved in iso-propanol for later polymerase chain reaction (PCR) analysis. For identification purposes, the larvae from each sampling location were placed in a plastic bucket leaving a breeding space on top and closed with a lid. Both buckets and lids were both marked to indicate the locality and date for the collection and then transported to NMCP's insectary, where the larvae were transferred by pipette into marked plastic bowls with clean water placed on a cage to secure them from being accidentally tilted and spilled while in the insectary. Each bowl was covered with a net and a cotton wool soaked in 10% sugar solution was placed on top for feeding. A heater was used to ensure maintenance of warm temperatures and to facilitate growth. When the larvae hatched and became adults, they were removed using a sucking tube, placed in marked paper cups and covered with a net. The mosquitoes were still fed by the sugar solution, which was placed on the side of the cups for daily morphological identification by the entomologist.

### Remotely sensed data

Mosquito vector habitat requires specific characteristics, such as sufficient surface water for reproduction, a certain humidity level for

adult mosquito survival and a suitable temperature allowing acceptable development rates for both vector and parasite (Ceccato *et al.*, 2005). For that reason, for a period starting 4 weeks prior to the study until the end of the entomological survey, Collection-5 (Land Surface, Temperature & Emissivity data) from the Moderate Resolution Imaging Spectroradiometer (MODIS) (<http://modis.gsfc.nasa.gov>) was downloaded from the Land Processes Distributed Active Archive Center in USA (<https://lpdaac.usgs.gov/>) for land surface temperature (LST) data at 1-km spatial resolution. We relied on the MOD11A2 product, which is the average value of clear-sky LSTs during an 8-day period. In addition, a 24-day temperature average was obtained for the three weeks prior to the entomological survey. Rainfall estimates (RFE) at 8-km spatial resolution were downloaded from the Famine Early Warning Systems Network (FEWS NET) Africa (<http://earlywarning.usgs.gov/afghan/downloads/index.php?regionID=af&productID=3&periodID=6>). In addition, a digital elevation model (DEM) from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (<http://asterweb.jpl.nasa.gov>) at 30-m spatial resolution was used for altitude data.

High-resolution data from the RapidEye satellite (<http://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/rapid->



**Figure 6.** Example of the final potential vector breeding sites classification of the malaria-endemic area covered by the 5 m resolution RapidEye data (northern part of the Lowveld).

eye/) obtained from the ESA Data Warehouse ([http://gmesdata.esa.int/web/gsc/news/latest\\_20110607](http://gmesdata.esa.int/web/gsc/news/latest_20110607)) covering the malaria-endemic area of Swaziland at 6.5-m spatial resolution was used in order to determine the various land cover types. The RapidEye pixel resolution was resampled to 5-m pixel size during ortho-rectification by the data provider. A standardized atmospheric correction was applied to the 62 RapidEye tiles obtained to ensure consistent image pre-processing and thus a more reliable land cover classification. Land use/land cover (LULC) classification of the high-resolution data was conducted by applying an object-based image analysis with a predefined hierarchical Ruleset (supervised classification) using the software eCognition (Trimble GeoSpatial, Munich, Germany). The overall accuracy of the land cover classification was 80.7% with a Kappa coefficient of 0.78. As separate layers, high-resolution, Euclidean distance-to-land cover layers were generated as input for the statistical modelling of potential mosquito breeding sites. Figure 2A shows the land cover classification, while Figure 2B-D shows a set of three examples of distance-to-land cover layers used in the statistical model for prediction. The high spatial resolution of the land cover classes allowed the identification of small water bodies and wetlands that can be potential breeding sites for malaria vectors. An analysis, based on a total of 18 different environmental variables for each larvae sampling point (Table 1), was carried out with the aim of identifying which parameters that best describe a vector breeding focus and its environmental requirements (Figures 3A-D and 4A-D). Eight LULC classes were used: settlements, subsistence farming, large-scale agriculture, savannah, forest, bush, bare soil/rocks and roads. All eight distance-to-land cover variables were categorized into three distance categories, which were defined by the 33<sup>rd</sup> and 66<sup>th</sup> centiles. This made it possible to estimate the effect of the different distances on the presence or absence of vector breeding sites. The distances from the vector breeding sites to each of the distance-to-land cover classes were defined by calculating the Euclidean distance from each sampling point to the centroid of the pixel with the specific land cover type.

### Statistical analysis

The relationship between the above environmental factors and the presence or absence of larvae were analyzed using a data mining approach in the statistical software See5 (RuleQuest Research, St Ives, Australia), which uses a Decision Tree R3 Ruleset induction engine developed by Quinlan (1993). The algorithm classifies data according to independent variables a branching method, which splits the data to illustrate every possible outcome of a probability-based decision. ArcGIS 10.1 (ESRI, Redlands, CA, USA) was used to conduct zonal statistical analyses at each larval sampling location in order to retrieve the environmental parameters for each location. The Decision Tree Ruleset was thus based both on larval field data from the sampling points and the local environmental variables.

A training dataset was used to select independent variables relevant for the classification of the presence of larvae and to obtain the Decision Tree Ruleset for classification accuracy using the field data provided. Due to the low sample size, the analysis was performed based on the results from all 21 positive sampling points. The classifier construction options in the statistical analysis were set so that it would provide a final Ruleset for classification of unknown water bodies/wetlands. The location of standing water bodies and wetlands from the high-resolution LULC map was afterwards applied to the final Ruleset for analysis. Based on the final classification rule, a map of all potential breeding sites in the malaria-endemic area of Swaziland was produced.

### Logistic regression

As a first step, the data were explored for any redundancies between variables via a statistical correlation analysis using Pearson's correlation coefficient. Bivariate analyses were performed to determine the relationship between each of the environmental covariates and presence of vector breeding sites using logistic regression (Stata Statistical Software, version 13.0). In order to evaluate and compare the results from the Decision Tree analysis a backward-selection, stepwise logistic regression analysis with a 15% significance level for removal was carried out based on the same input data set as used in the Decision Tree approach. Covariates included in the stepwise regression were all those significant at the 15% significance level in the bivariate results. We used the 15% significance level to reduce the chance of excluding important predictors from the stepwise multiple regression. This analysis enabled finding the most effective and parsimonious set of variables predicting the dependent variable (presence or absence of larvae).

## Results and Discussion

### The field campaign

Larvae were found in 21 out of the total of 30 sample sites (70%) of water bodies. However, these were mostly areas where it had not rained in the past two to three weeks prior to sampling. In areas where it had rained recently, no larvae were found and it was difficult to ascertain whether this was due to the fact that the site was not suitable or the larvae had been washed away due to the rain.

In total, 152 mosquito larvae were collected during the field scooping campaign. Upon rearing, about 60% of the captured larvae was found to be *An. pretoriensis* and 15 % *An. gambiae*. The rest of the mosquitoes were *Culicines* (20%) and *Culex* (5%) species. Sites where no larvae were found were mainly large dams, open turbulent dams and areas where rain was reported less than a week prior to sampling. Most of the larvae were collected from the edges of shallow and sheltered water bodies, but where there was enough sunlight. Larvae were also found in clear pools of water with little aquatic flora and fauna. Geographically, the highest larval densities were found in the southern and northern part of the Lowveld as well as on the eastern plateau bordering Mozambique. These were areas with large-scale irrigated farms such as sugarcane surrounded by subsistence farming. Excess runoff water was common in these areas as most of the fields had sprinklers continually wetting various parts of the field; hence drainage water accumulated in small depressions around each field assuring availability of water also in the dry season.

### Decision Tree analysis

The analysis of the 21 sample points with larvae using 18 environmental variables showed that the combined use of only 3 variables in 3 rules correctly classified 95.2% of the larvae present sites and they included: distance-to-subsistence farming, 10-day RFE of the second 10-day period and distance-to-savannah. The variable with the highest attribute usage was the distance-to-subsistence farming that appeared in 90% of the rulesets in predicting a class. Indeed, distance-to-subsistence farming less than 216 m turned out to be a major determinant for vector breeding sites (Figure 5). However, no relation could be found between breeding sites and distance-to-settlements. In addition to the variable distance-to-subsistence farming, the 10-day RFE with 1-week lag had an attribute usage of 38% in the ruleset and the distance-to-



grassland/savannah showed an attribute usage of 33%. According to these results, the LST over the 4 weeks before the sampling and altitude had no influence on the occurrence of vector breeding sites. This was also the case for all other LULCs except distance-to-subsistence farming and grassland/savannah and the other three remaining RFEs. As a final step, the classification ruleset derived from the Decision Tree analysis, which used the variables distance-to-subsistence farming, distance-to-grassland/savannah and the 10-day RFE, was applied to the wetland and water body layer with 5-m spatial resolution in order to identify potential vector breeding sites with similar environmental characteristics as those identified through the entomological survey. The resulting map of predicted vector breeding sites for the whole malaria-endemic area in Swaziland identified potential breeding sites of various sizes (from 5 to >100 m of standing water). This high-resolution map can be used as a guide for larval sampling activities. Figure 6 shows an example of the final potential vector breeding sites classification of the malaria area covered by the 5-m resolution RapidEye data (northern part of the Lowveld) in Swaziland.

### Logistic regression

The exploration for variable redundancies via correlation analysis showed that the average temperature for the 24 days before the larval sampling, which is an aggregate measure of all four temperature periods, was highly correlated with all other temperature variables (Pearson's  $r > 0.8$ ) and was therefore removed from the analysis. The bivariate regression analysis indicated that the following variables were significant at the 15% significance level: distance-to-subsistence farming, temperature and rainfall of the fourth week prior to larval sampling, altitude and distance-to-settlement. Results from the step-wise logistic regression suggest that the best model was the one including only distance-to-subsistence farming. The higher the distance from subsistence farming the lower the odds of presence of mosquito breeding sites with an odds ratio (OR) 0.21 [95% confidence interval (CI): 0.06-0.71] and a P value of 0.012 (Table 2).

### Conclusions

In this study, we argue that exclusive reliance on RS and knowledge on how climatic factors and other environmental variables influence vector breeding is inadequate for the prediction of potential breeding sites, unless conducted with satellite-generated, high-resolution imagery in conjunction with information on presence or absence of larvae. To effectively guide control activities, it is imperative that actual entomological data on vector breeding and distribution be taken into account, e.g., when constructing regression models, so that previously not fully understood ecological effects and spatial breeding site heterogeneity can be elucidated. Linking entomological field collection with adequate remotely sensed environmental data does not only increase the accuracy of prediction models, but it also assists specification and identification of spatial heterogeneities with regard to vector breeding habitats. Understanding such variations may further facilitate determining the contribution and impact of vector control and other malaria control measures. In addition, as countries move towards malaria elimination, endemic transmission becomes limited to residual foci (Cohen *et al.*, 2013), where prevention of larval breeding will be necessary. Most countries already have geographic databases on water bodies as well as entomological data such as larva scooping. It would be important that this information is incorporated into future databases.

Due to the current availability of only limited number of studies with detailed and explicit spatio-temporal variations on malaria transmis-

sion in Swaziland, it is not surprising that the country has not yet reached the goal of zero malaria cases, although current control efforts have been successful to reduce the overall burden. The present work is the first attempt to map potential breeding sites in Swaziland using remotely sensed data in conjunction with information on presence/absence of mosquito larvae. We have demonstrated that distance-to-subsistence farming is the main predictor for presence of active breeding sites. This can be explained by the fact that subsistence farming is a full-time occupation for most rural communities and the proximity of human hosts may explain the strong association between distance-to-subsistence farming and vector breeding sites. This finding is also consistent with a study by Ahmad *et al.* (2011), which deals with vector breeding habitats located between 100 and 400 m from human settlements. In the rural areas of Swaziland, almost every household is surrounded by hectare-sized subsistence farming fields and this translates to a continuous land use type as fields join at the edges. This situation supports the presence of foci and is thus an important factor for malaria endemicity. Similar result was found in a study by Li *et al.* (2008), which concluded that houses in great proximity to streams have more abundant mosquitoes than other places and that breeding was high in the nearby valley bottoms.

No relation could be found between breeding sites and distance-to-settlements in the present study. This could further be explained that cattle are not near settlements but exist in abundance at subsistence farming sites. From the field survey results, it was clear that the vector breeds very well in small water bodies down to the size of cattle hoof prints, which provide sheltered conditions. Turbulent water does not favour larvae development, which has already been observed in previous studies (e.g. Ageep *et al.*, 2009), while Dejenie *et al.* (2011) remark that it is very likely that peripheral water bodies from irrigation channels and small depressions like cattle hoof prints, rather than the larger main dams, become preferred sites for mosquito larvae. Hence, it is very likely that locations characterized by cattle and proximity to water bodies are at risk due to the probable presence of vector breeding sites.

Rainfall was not found to be associated with vector breeding, which is not surprising as the more frequent and heavy the rainfall, the higher the possibility that mosquito larvae are washed away (Savage *et al.*, 1990). As already alluded to, mosquito larvae were not likely to be found in areas reporting rain less than a week prior to the survey. However, as irrigation played a major role in providing excess water, it was not unexpected that rainfall had a lower attribute usage percentage (38%) in the Decision Tree analysis. Indeed, availability of water also in the dry season results in suitable habitats the year round facilitating oviposition compared to other parts where irrigation is not applied. This tells us that rainfall may not always be a useful predictor and that mosquitoes can be plentiful even without rainfall. Based on the discussion above, it would not be unusual that the use of rainfall as a covariate in logistic regression models even indicates a negative correlation, which is not always the case with other studies analyzing environmental factors in relation to malaria transmission (Briët *et al.*, 2008; Laneri *et al.*, 2010). Thus, in areas where larval breeding occurs in the absence of rainfall, other environmental proxies will have to be used to identify and map potential breeding sites.

The maps produced in this study are not only useful for targeting residual foci, but could also help reducing the LSM budget. This would be achieved through the implementation of larvicide application guided by risk maps (Dambach *et al.*, 2014).

**Table 1. Data sources and properties of the environmental covariates used in predicting potential mosquito breeding sites.**

Type of data	Source	Date	Temporal resolution (days)	Spatial resolution
Altitude	DEM	2012	-	30 m
Land cover	RapidEye	2011	-	5 m
LST	MODIS data	01.11.2012-05.11.2012	8	1 km
RF	FEWS NET	01.11.2012-10.12.2012	10	8 km

DEM, digital elevation model (<http://asterweb.jpl.nasa.gov/>); RapidEye, RapidEye satellite imagery ([http://gmesdata.esa.int/web/gsc/news/latest\\_20110607/](http://gmesdata.esa.int/web/gsc/news/latest_20110607/)); LST, land surface temperature; MODIS, Moderate Resolution Imaging Spectroradiometer (<http://modis.gsfc.nasa.gov/>); RF, rainfall; FEWS NET, Famine Early Warning Systems Network (<http://earlywarning.usgs.gov/afghan/downloads/index.php?regionID=af&productID=3&periodID=6>).

**Table 2. Analytical results using three different statistical methodologies.**

Predictor	Bivariate logistic regression		Stepwise logistic regression		Decision Tree analysis	
	OR (95% CI)	P <sup>o</sup>	OR (95% CI)	P	Ruleset attribute usage (%)	Cumulative classification accuracy (%)
<b>Subsistence farming<sup>#</sup></b>						
Distance (m): <370	1.00	-	-	-	-	-
Distance (m): 370-470	0.05 (0.01; 0.49)	0.003	0.05 (0.01; 0.49)	0.009	90	47.6
Distance (m): ≥470 m	0.21 (0.06; 0.71)	0.003	0.21 (0.06; 0.71)	0.012	90	47.6
<b>Distance-to-savannah (m)</b>						
<150	1.00	-	-	-	-	-
150-620	0.56 (0.90; 3.52)	0.536	-	33	95.2 <sup>§</sup>	-
≥620	0.93 (0.38; 2.24)	0.864	-	33	95.2 <sup>§</sup>	-
<b>Temperature</b>						
First week	0.64 (0.25; 1.63)	0.337	-	-	-	-
Second week	0.76 (0.31; 1.87)	0.549	-	-	-	-
Third week	0.93 (0.39; 2.25)	0.873	-	-	-	-
Fourth week <sup>#</sup>	3.98 (1.20; 13.23)	0.008	-	-	-	-
24-day average temperature	0.75 (0.30; 1.84)	0.520	-	-	-	-
<b>Rainfall</b>						
First week	0.80 (0.33; 1.95)	0.614	-	-	-	-
Second week	1.03 (0.43; 2.48)	0.951	-	38	76.2 <sup>^</sup>	-
Third week	0.81 (0.34; 1.97)	0.647	-	-	-	-
Fourth week <sup>#</sup>	2.57 (0.85; 7.79)	0.060	-	-	-	-
Altitude <sup>#</sup>	3.50 (1.08; 11.29)	0.016	-	-	-	-
<b>Distance-to-bare soil/rocks (m)</b>						
120	1.00	-	-	-	-	-
120-700	1.20 (0.22; 6.68)	0.835	-	-	-	-
≥700	0.75 (0.30; 1.88)	0.536	-	-	-	-
<b>Distance-to-bush land (m)</b>						
<420	-	-	-	-	-	-
420-460	1.25 (0.22; 7.08)	0.801	-	-	-	-
≥460	0.94 (0.36; 2.41)	0.890	-	-	-	-
<b>Distance-to-forest (m)</b>						
<700	1.00	-	-	-	-	-
700-720	1.75 (0.31; 10.02)	0.528	-	-	-	-
≥720	0.72 (0.26; 2.00)	0.525	-	-	-	-
<b>Distance-to-agriculture (m)</b>						
<200	1.00	-	-	-	-	-
200-450	2.63 (0.45; 15.31)	0.277	-	-	-	-
≥450	1.39 (0.57; 3.43)	0.465	-	-	-	-
<b>Distance-to-main road (m)</b>						
<340	-	-	-	-	-	-
340-720	1.80 (0.32; 10.20)	0.504	-	-	-	-
≥720	1.16 (0.46; 2.88)	0.757	-	-	-	-
<b>Distance-to-settlement<sup>#</sup> (m)</b>						
<100	-	-	-	-	-	-
100-510	0.38 (0.07; 2.22)	0.277	-	-	-	-
≥510	0.39 (0.12; 1.30)	0.089	-	-	-	-

OR, odds ratio; CI, confidence interval. <sup>o</sup>Based on likelihood ratio test; <sup>#</sup>variable included in the stepwise regression; <sup>§</sup>combined use of subsistence farming, rainfall in the 2nd, week and savannah; <sup>^</sup>combined use of subsistence farming and rainfall\_week\_2.



## References

- Ageep TB, Cox J, Hassan MM, Knols BGJ, Benedict MQ, Malcolm CA, Babiker A, El Sayed BB, 2009. Spatial and temporal distribution of the malaria mosquito *Anopheles arabiensis* in northern Sudan: influence of environmental factors and implications for vector control. *Malaria J* 8:123.
- Ahmad R, Ali WN, Nor ZM, Ismail Z, Hadi AA, Ibrahim MN, Lim LH, 2011. Mapping of mosquito breeding sites in malaria endemic areas in Pos Lenjang, Kuala Lipis, Pahang, Malaysia. *Malaria J* 10:361.
- Beck LR, Lobitz BM, Wood BL, 2000. Remote sensing and human health: new sensors and new opportunities. *Emerg Infect Dis* 6:217-27.
- Bøgh C, Lindsay SW, Clarke SE, Dean A, Jawara M, Pinder, M, Thomas CJ, 2007. High spatial resolution mapping of malaria transmission risk in the Gambia, west Africa, using LANDSAT TM satellite imagery. *Am J Trop Med Hyg* 76:875-81.
- Briët OJ, Vounatsou P, Gunawardena DM, Galappaththy GN, Amerasinghe PH, 2008. Temporal correlation between malaria and rainfall in Sri Lanka. *Malaria J* 7:77.
- Ceccato P, Connor SJ, Jeanne I, Thomson MC, 2005. Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. *Parassitologia* 47:81-96.
- Chanda E, Doggale C, Pasquale H, Azairwe R, Baba S, Mnzava A, 2013. Addressing malaria vector control challenges in South Sudan: proposed recommendations. *Malaria J* 12:59.
- Clennon JA, Kamanga A, Musapa M, Shiff C, Glass GE, 2010. Identifying malaria vector breeding habitats with remote sensing data and terrain-based landscape indices in Zambia. *Int J Health Geogr* 9:58.
- Cohen JM, Dlamini S, Novotny JM, Kandula D, Kunene S, Tatem AJ, 2013. Rapid case-based mapping of seasonal malaria transmission risk for strategic elimination planning in Swaziland. *Malaria J* 12:61.
- Dambach P, Sié A, Lacaux J-P, Vignolles C, Machault V, Sauerborn R, 2009. Using high spatial resolution remote sensing for risk mapping of malaria occurrence in the Nouna district, Burkina Faso. Available from: <http://www.globalhealthaction.net/index.php/gha/article/view/2094>
- Dambach P, Traoré I, Becker N, Kaiser A, Sié A, Sauerborn R, 2014. EMIRA: ecologic malaria reduction for Africa; innovative tools for integrated malaria control. Available from: <http://www.globalhealthaction.net/index.php/gha/article/view/25908>
- Dejenie T, Yohannes M, Assmelash T, 2011. Characterization of mosquito breeding sites in and in the vicinity of tigray microdams. *Ethiop J Health Sci* 21:57-66.
- Kulkarni MA, Desrochers RE, Kerr JT, 2010. High resolution niche models of malaria vectors in northern Tanzania: a new capacity to predict malaria risk? *PloS One* 5:e9396.
- Laneri K, Bhadra A, Ionides EL, Bouma M, Dhiman RC, Yadav RS, Pascual M, 2010. Forcing versus feedback: epidemic malaria and monsoon rains in Northwest India. *PLoS Comput Biol* 6:e1000898.
- Li L, Bian L, Yan G, 2008. A study of the distribution and abundance of the adult malaria vector in western Kenya highlands. *Int J Health Geogr* 7:50.
- Malaria Indicator Survey, 2010. Swaziland 2010. Available from: <http://malariasurveys.org/surveys.cfm?country=Swaziland%202010#5010>
- Oesterholt MJAM, Bousema JT, Mwerinde OK, Harris C, Lushino P, Masokoto A, Mwerinde H, Mosha FW, Drakeley CJ, 2006. Spatial and temporal variation in malaria transmission in a low endemicity area in northern Tanzania. *Malaria J* 5:98.
- Quinlan JR, 1993. C4.5: programs for machine learning. Morgan Kaufmann Publ., San Francisco, CA, USA.
- Savage HM, Rejmankova E, Arredondo-Jimenez JI, Roberts DR, Rodr'iguez MH, 1990. Limnological and botanical characterization of larval habitats for two primary malarial vectors, *Anopheles albimanus* and *Anopheles pseudopunctipennis*, in coastal areas of Chiapas State, Mexico. *J Am Mosquito Contr* 6:612-20.