Earth observation in support of malaria control and epidemiology: MALAREO monitoring approaches

Jonas Franke,[1] Michael Gebreslasie,[2] Ides Bauwens,[3] Julie Deleu,[4] Florian Siegert[1,5]

[1]Remote Sensing Solutions GmbH, Baierbrunn, Germany; [2]School of Agricultural, Earth and Environmental Sciences, University KwaZulu-Natal, Durban, South Africa; [3]Nazka mapps bvba, Hever; [4]EUROSENSE, Wemmel, Belgium; [5]Biology Department II, GeoBioCenter, Ludwig-Maximilians-University, Planegg, Germany

Abstract

Malaria affects about half of the world's population, with the vast majority of cases occuring in Africa. National malaria control programmes aim to reduce the burden of malaria and its negative, socio-economic effects by using various control strategies (*e.g.* vector control, environmental management and case tracking). Vector control is the most effective transmission prevention strategy, while environmental factors are the key parameters affecting transmission. Geographic information systems (GIS), earth observation (EO) and spatial modelling are increasingly being recognised as valuable tools for effective management and malaria vector control. Issues previously inhibiting the use of EO in epidemiology and malaria control such as poor satellite sensor performance, high costs and long turnaround times, have since been resolved through modern technology. The core goal of this study was to develop and implement the capabilities of EO data for national malaria control programmes in South Africa, Swaziland and Mozambique. High- and very high resolution (HR and VHR) land cover and wetland maps were generated for the identification of potential vector habitats and human activities, as well as geo-information on distance to wetlands for malaria risk modelling, population density maps, habitat foci maps and VHR household maps. These products were further used for modelling malaria incidence and the analysis of environmental factors that favour vector breeding. Geo-products were also transferred to the staff of national malaria control programmes in seven African countries to demonstrate how EO data and GIS can support vector control strategy planning and monitoring. The transferred EO products support better epidemiological understanding of environmental factors related to malaria transmission, and allow for spatio-temporal targeting of malaria control interventions, thereby improving the cost-effectiveness of interventions.

Introduction

There is a vital need for new strategies and innovative health-related monitoring approaches that support the reduction of the burden of environmentally related diseases such as malaria and related negative, socio-economic effects. According to the World Health Organization (WHO), malaria is currently endemic in more than 100 countries worldwide, affects about half of the world's population, and sees the largest number of cases reported in Africa (WHO, 2013b). Each year, approximately 220 million cases and 650 thousands deaths are reported (WHO, 2012b). In Africa, the highly efficient *Anopheles gambiae* mosquito is the dominant vector of *Plasmodium falciparum*, the most dangerous of the four human malaria parasites. Besides the serious situation it poses for the infected individual, malaria has far-reaching impacts on both the private and the public healthcare sector, which ultimately impede the economic development of countries or regions. In addition, increased travel and migration have caused many imported cases outside of endemic areas. For instance, over 100,000 imported malaria cases have been reported in Europe in the last 10 years by travellers and immigrants, and even local transmission from imported cases has been reported (WHO, 2012a). Malaria is influenced strongly

by the environment, since the mosquito vector requires specific habitats with sufficient surface water for reproduction, a certain humidity level for adult mosquito survival and temperatures which influence the development rates of both the vector and the parasite (Ceccato *et al.*, 2005). The area affected by and intensity of malaria depend on vector abundance, the control of which is a basic technical element of the global disease control strategy (WHO, 2013a).

Several factors that render an area attractive to mosquitoes, such as proximity to water bodies, land use and land cover (LULC), house structures and even human behavioural patterns, can be derived from remote sensing applications. Earth observation (EO) can be used for the surveillance, monitoring and early warning of diseases closely associated with environmental conditions. Remotely sensed images holding information on climatic conditions are powerful predictors of vector distribution patterns and their average malaria parasite transmission level (Rogers *et al.*, 2002). In epidemiology, the link between physical measurements of reflected radiation and measures of a disease, vector or intermediate host, has already been established via logical sequences (Crombie *et al.*, 1999; Curran *et al.*, 2000). Remote sensing has been used to associate LULC types with vector habitats based on simple classification techniques, as well as complex statistical models linking satellite-derived meteorological and EO data with vector biology and abundance (Kalluri *et al.*, 2007; Guerra *et al.*, 2008). Environmental factors such as land cover, vegetation green-up, permanent water, wetlands, soil moisture and human settlements have a strong influence on malaria patterns (Beck *et al.*, 2000; Tatem *et al.*, 2004; Zeilhofer *et al.*, 2007; Dambach *et al.*, 2009).

Interdisciplinary research over the last decades resulted in a better understanding of the social and economic dimensions of malaria and facilitated new research into and development of tools leading to advances in vector control. Building on this knowledge, there is now a vital need for new monitoring capabilities that directly support disease control and reduce the burden of environmental related diseases in vulnerable countries. An area where *Anopheles* sustains the basic reproductive rate required for malaria transmission is called an active focus. Areas within the focus that exceed the average transmission intensity are known as hotspots. According to Bousema *et al.* (2012), the identification and targeting of such foci and hotspots form the cornerstone of successful malaria control and elimination, while disease modelling – combining environmental data with population data and malaria case data – is excellent for defining malaria transmission foci. However, the spatial scale of these foci is often not sufficient to detect malaria transmission hotspots (Bejon *et al.*, 2010, Bousema *et al.*, 2010). In the past, the suitability of EO data in spatial and temporal disease pattern studies was restricted by the data's low spatial resolution, high costs and long turnaround times for products. Beck *et al.* (2000) indicate that new sensors fulfill the spatial and temporal requirements needed to make remote sensing a functional tool in disease surveillance and control. Ceccato *et al.* (2002) reviewed the capabilities of remote sensing and later stated that *The time is ripe for the wealth of research knowledge and products from developed countries be made available to the decision-makers in malarious regions of the globe where this information is urgently needed* (Ceccato *et al.*, 2005).

A primary objective of the present study was to explore the potential of current satellites and to add value to existing disease-related EO approaches by using high-resolution (HR) and very high resolution (VHR) imagery not used thus far for vector control and epidemiological studies. MALAREO, a research project partly funded by the European Commission under the seventh framework programme (FP7), aimed to assess the benefits of new EO monitoring capabilities, enhancing disease control strategies, reducing disease control costs, providing reliable decision support information and supporting sustainable disease elimination. In cooperation with the national malaria control programmes (NMCPs) of South Africa, Swaziland and Mozambique, EO-based geo-information products have been developed that directly support malaria control interventions and epidemiological research in these countries. The MALAREO approach made use of HR and VHR satellite sensors for generating geo-products with the spatial and thematic detail required for detailed and reliable decision support. This article describes the technical details of the developed EO-based geo-information and the methodologies applied to generate the EO-products that support the work of the NMCPs (MALAREO partner) in southern Africa.

Materials and Methods

Study area

The MALAREO study area (about 40,000 km^2) is located in southern Africa and covers southern Mozambique, eastern Swaziland and the part of KwaZulu-Natal affected by malaria in South Africa (Figure 1). It is located at the fringe of the malaria transmission area and contains endemic and epidemic malaria areas. The three countries are in different stages of malaria elimination and have therefore different requirements regarding monitoring disease risk and its related environmental factors. Cross-border migration causes a considerable amount of imported cases that can result in new outbreaks and local cases. An important influence in the study area is the Lubombo Spatial Development Initiative (LSDI), which includes a cross-border malaria control program, jointly implemented by Mozambique, South Africa and Swaziland. A significant reduction in parasite prevalence among the human population, vector density and sporozoite positivity rates have been documented following the launch of the comprehensive malaria control program (Sharp *et al.*, 2007). Recently, funding by LSDI was suspended in Mozambique, which will likely affect its malaria control substantially.

Mozambique

Malaria is a major cause of morbidity and mortality in Mozambique. According to the Ministry of Health, estimated prevalence in two- to nine-year-olds varies from 40 to 80% (MISAU, 2011). Malaria is endemic and climate favours its transmission throughout the year, with a peak after the rainy season. The NMCP, with its current remit, was established in 1982. Mozambique is currently at the malaria-control stage.

South Africa

Regions affected by malaria include the eastern parts of Limpopo and Mpumalanga province, and the northeastern part of KwaZulu-Natal. Malaria transmission is spatially variable and seasonal, peaking in the warm and rainy summer months. The South African Government and the Southern African Development Community (SADC) have supported intensive malaria control activities over many decades (Balfour, 2002), and succeeded in halting transmission in most of the country, but malaria remains endemic in the study area. South Africa is currently at the malaria pre-elimination stage and continued support is required to steer its malaria elimination objectives (Moonasar *et al.*, 2012).

Swaziland

Swaziland's national malaria burden has been greatly reduced, with laboratory-confirmed cases decreasing from 3.9 to 0.07 cases per 1000 individuals between 1999 and 2009 (Kunene *et al.*, 2011). This reduction has been attributed to the recent scale-up of vector control activities in the at-risk region and in areas associated with the cross-border collaboration with Mozambique and South Africa. Recognizing Swaziland's success, the SADC and the African Union forecasted that Swaziland will reach malaria elimination by 2015 (African Union, 2007). Swaziland is currently at the malaria elimination stage. A location-based malaria case tracking system, employing global positioning systems (GPS), was introduced by the NMCP and has been active since 2010.

Data

Many HR sensors can be used to monitor LULC change, for example, to discriminate perennial and transient water bodies and wetlands or identify human activities. HR instruments such as RapidEye (http://blackbridge.com/rapideye/), SPOT (http://www.geo-airbusds.

com/en/143-spot-satellite-imagery) and FORMOSAT-2 (http://www.geo-airbusds.com/en/160-formosat-2) have the advantage of a wide swath width and thus large area coverage. In the context of house detection when identifying human settlements, VHR instruments allow feature mapping in great spatial detail. However, these VHR data are cost-intensive, limiting their suitability for malaria control. MALAREO focused on three different types of EO applications, namely, EO for land cover mapping, EO to directly support the work of the NMCPs and EO for epidemiological studies. Each of these applications have specific data characteristic requirements, as well as spatial coverage and data analysis techniques. The data used in this study are described below by application.

Earth observation data used for land cover mapping

For land cover mapping, data of the RapidEye satellites were used. This is a constellation of five identical satellites that collect data in 5 spectral bands, and was launched in August 2008. The sensors' ability to acquire data from several viewing angles allows imagery to be captured at one location on a daily basis (5.5 days at nadir) (Tyc *et al.*, 2005). The pixel resolution is 6.5 m, which the data provider resamples
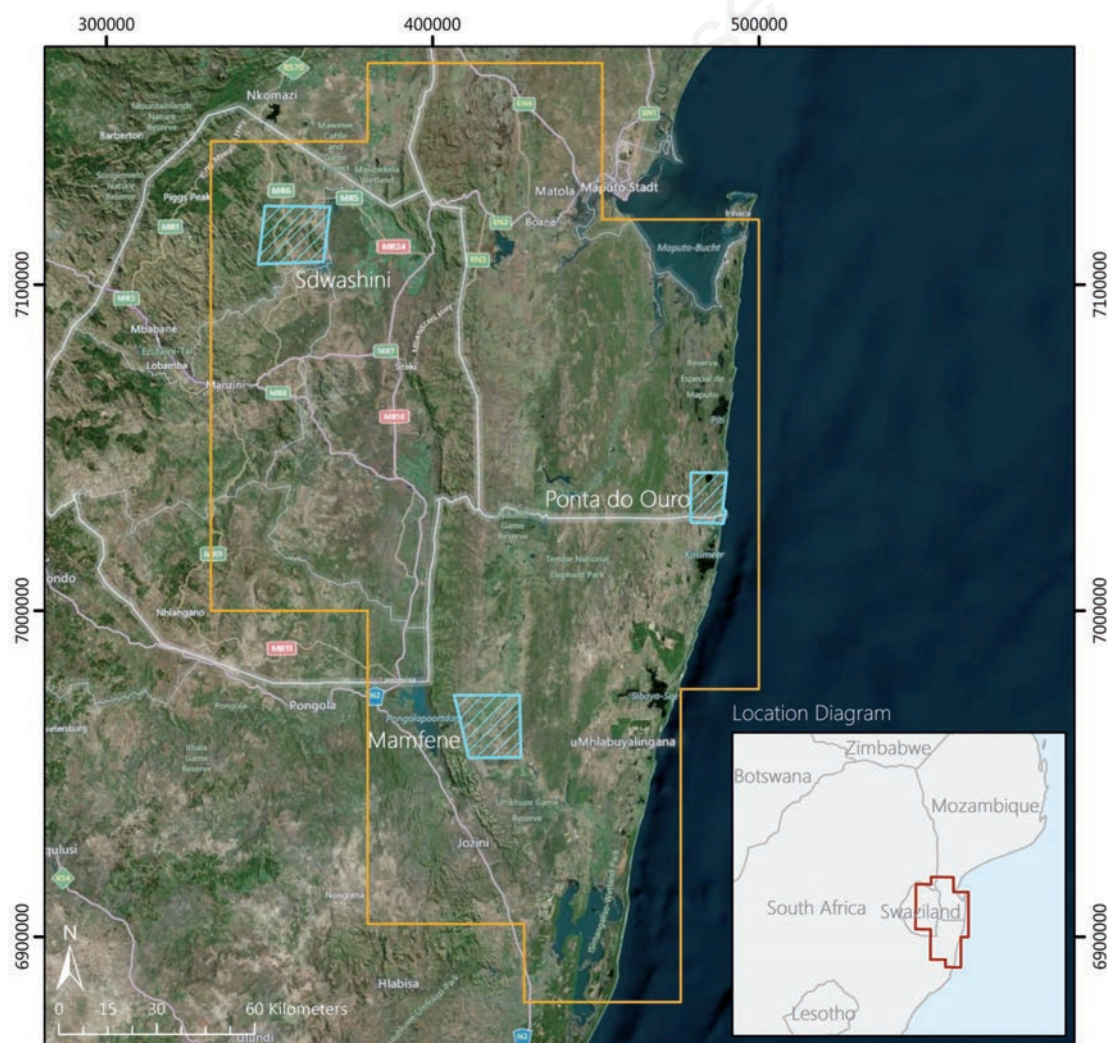

Figure 1. MALAREO study area with high resolution (yellow) and very high resolution (blue) test sites. Backgroud: BingMaps.

to 5 meters during orthorectification. About 35,000 km^2 orthorectified RapidEye archive data (62 25x25 km tiles) were ordered via the Copernicus Space Component Data Access system (CSCDA) (https://copernicusdata.esa.int/) (Figure 2). The data were acquired in 2011 on 5 different dates (between 18 July and 10 November). Only one of the 62 tiles was acquired on 18/07/2011, while 61 tiles were acquired within a short period at the beginning of the rainy season, *i.e.* between the end of October and mid-November 2011. Hence, seasonal differences between the scenes were negligible.

Earth observation data for products supporting the national malaria control programmes

These products were developed with the aim of directly supporting NMCP managements. In addition to land cover information generated from HR data, VHR images can help to identify houses and small huts. The only already available VHR image sources are free image services such as Google Earth, Microsoft BingMaps or ESRI's World Imagery (ArcGIS Basemap Service). These were tested as an alternative to the cost-intensive VHR data for house mapping. Since additional VHR imagery was required for mapping houses and small huts, 850 km^2 VHR GeoEye-1 and IKONOS-2 (https://www.digitalglobe.com/) data were ordered via the ESA CSCDA. VHR imagery is cost-intensive, thus, only hotspot areas could be considered to be covered with VHR data for demonstration purposes. Three sites (Figure 1) of high relevance for the local malaria control activities were chosen to be assessed with VHR data: i) Mamfene in the Jozini local municipality, KwaZulu Natal, South Africa (covered by GeoEye-1 data; 0.5 m resolution in panchromatic); ii) Ponta do Ouro in the district of Matutuine, Maputo province, Mozambique (covered by IKONOS-2 data; 0.8 m spatial resolution in panchromatic); iii) Dwashini in the Hhohho District in Swaziland where an unexpected malaria outbreak occurred in April 2011 (covered by GeoEye-1 data; 0.5 m resolution in panchromatic).

Earth observation data used for malaria epidemiological studies

Low-resolution EO data were used to monitor environmental parameters as input for malaria incidence modelling. In this study, we used rainfall estimates (RFE) (temporal resolution: 10 days, spatial resolution: 8000 m) from the Famine Early Warning Systems Network (FEWS NET) (http://www.fews.net/) of the United States Agency for International Development (USAID). Land surface temperature & emissivity (temporal resolution: daily, spatial resolution: 1000 m), land cover (temporal resolution: yearly, spatial resolution: 500 m) and vege-
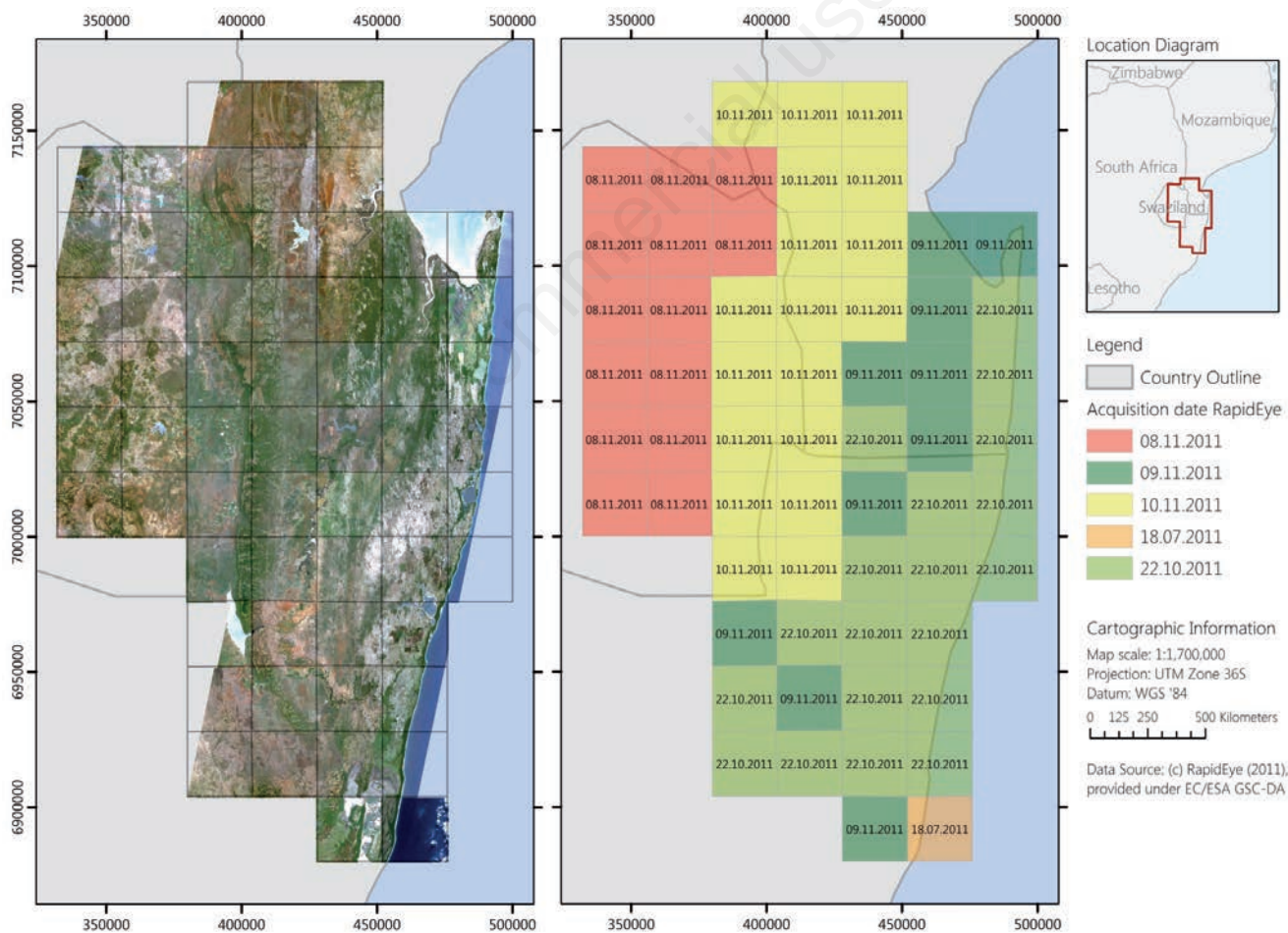

Figure 2. RapidEye image tiles (quicklooks) used in the study. RapidEye image (left) and the acquisition date per image tile (right). ©RapidEye (2011), provided under European Commission/European Space Agency Global Monitoring for Environment and Security Space Component Data Access.

tation indices (temporal resolution: 16 days, spatial resolution: 250 m) were obtained from the moderate resolution imaging spectroradiometer (MODIS) (http://modis.gsfc.nasa.gov/). The digital elevation model (DEM) (spatial resolution: 30 m) used came from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (http://asterweb.jpl.nasa.gov/gdem.asp). All these low resolution data have the benefit of being free of charge with good spatial and temporal coverage. The disadvantage is their low spatial resolution.

Methodology

Figure 3 shows MALAREO's EO-based, geo-information product development concept. Based on the NMCP's user requirements surveyed at the beginning of the project, EO-based, geo-information products of high relevance for improved planning of integrated vector control were identified. An EO feasibility study was realized afterwards in order to define required input data, map products and required method-

ologies to be applied in order to fulfil these user requirements. Some of these EO-based geo-information products were used as intermediate inputs in conjunction with additional malaria cases or entomological data for the malaria incidence modelling as well as for statistical vector breeding site analysis. The following chapters describe the underlying methodologies for each EO product shown in Figure 3.

Household mapping using very high resolution data

The integration of terrestrial geo-data into a computerised malaria control management system for spatio-temporal monitoring of indoor residual spraying (IRS) coverage, insecticide consumption and application rates has already been demonstrated in Mpumalanga, South Africa (Booman *et al.*, 2003). MALAREO aimed to expand this trend of geo-data use by adding an EO component for malaria control. VHR images from GeoEye-1 and IKONOS-2 of the three demonstration sites (Figure 1) were used to identify houses and small huts and to map their distri-
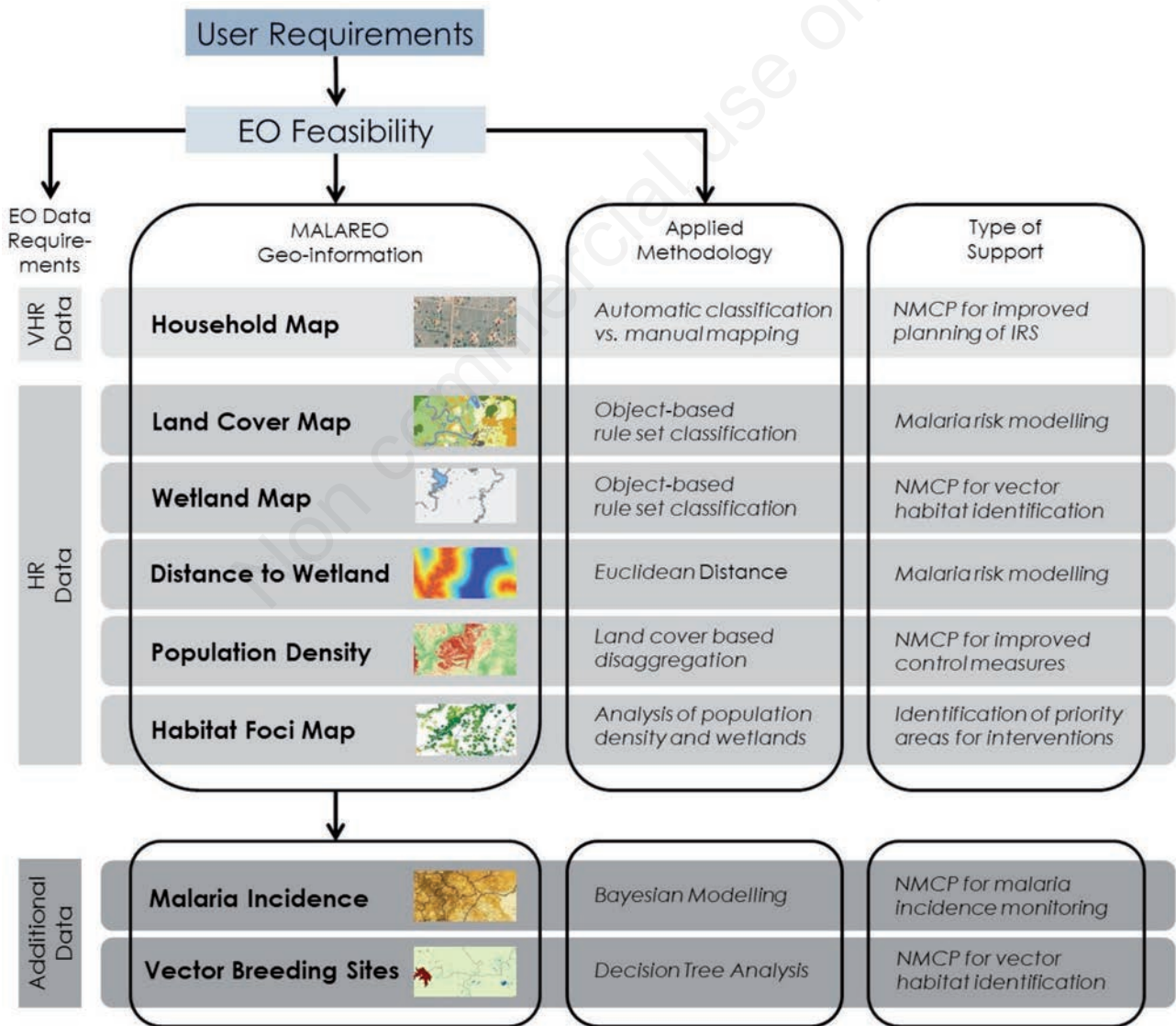

Figure 3. Concept of earth observation-based geo-information product development in MALAREO.

bution in these rural areas. Information on the structure type, their amount and distribution, were deemed valuable for a more efficient IRS, which is the main vector control intervention in the study area (together with distribution of insecticide-treated bed nets). National malaria control programmes aim to reduce the indoor parasite transmission rate via IRS of used houses and huts at the beginning of the wet season. The IRS campaigns target full area coverage (even remote single huts) in order to minimize all sources of malaria transmission. Since the suitability of applied insecticides depends on the type of building, geo-information on building structure, type and distribution is helpful. In the study area, the form of the buildings is an indication for the building material used, whereby round structures are mostly clay huts with thatched roofs, and square buildings are often made of bricks with roofs of corrugated tin. With such spatial information on structure distribution, the amount and type of insecticides needed per area can be better estimated and the progress and status of IRS coverage by the trained spraying teams can be monitored.

Mapping of houses and huts by remote sensing is challenging in rural areas, since most of the structures are very small with varying shape (round or rectangular) and numerous roof types that are hard to distinguish from other material (*e.g.* reed roofs). Taubenböck *et al.* (2010) presented an approach for feature detection using VHR data in urban areas. However, a review of the literature showed that no image analysis approach exists, that is feasible for the detection of small houses and huts in rural Africa. Figure 4 shows an example of a VHR

panchromatic GeoEye subset representing a homestead with various building types (round and rectangular) and demonstrating the very similar appearance of trees.

In this study, an approach for automated house/hut detection in rural African areas basing on VHR data was developed and tested. An object-based image classification using geometric, topological and spectral features was applied. As a first step, the multispectral bands of the VHR images were resampled to the spatial resolution of the panchromatic (black and white) band. The images were afterwards segmented into image objects by using all bands in eCognition software (Trimble Geospatial Imaging, Munich, Germany). The image is thereby segmented into objects of spatially adjacent and spectrally similar groups of pixels. The segmentation parameters (scale, shape and compactness of image objects) were set in a way that the image features of interest (round and squared houses) were optimally matched and could be differentiated from objects of other non-relevant land cover features. A ruleset was finally defined using geometric parameters such as size and compactness of the objects, topological parameters such as distance to other image objects, spectral parameters such as brightness of the objects, the normalized difference vegetation index (NDVI) as well as thematic information on land cover in order to reduce misclassifications (see below). The NDVI was used in order to differentiate between reed roofs and trees which often showed similar shapes of image objects (Figures 4 and 5). The final classifications of the three VHR images showed the spatial distribution of rectangular as well as round
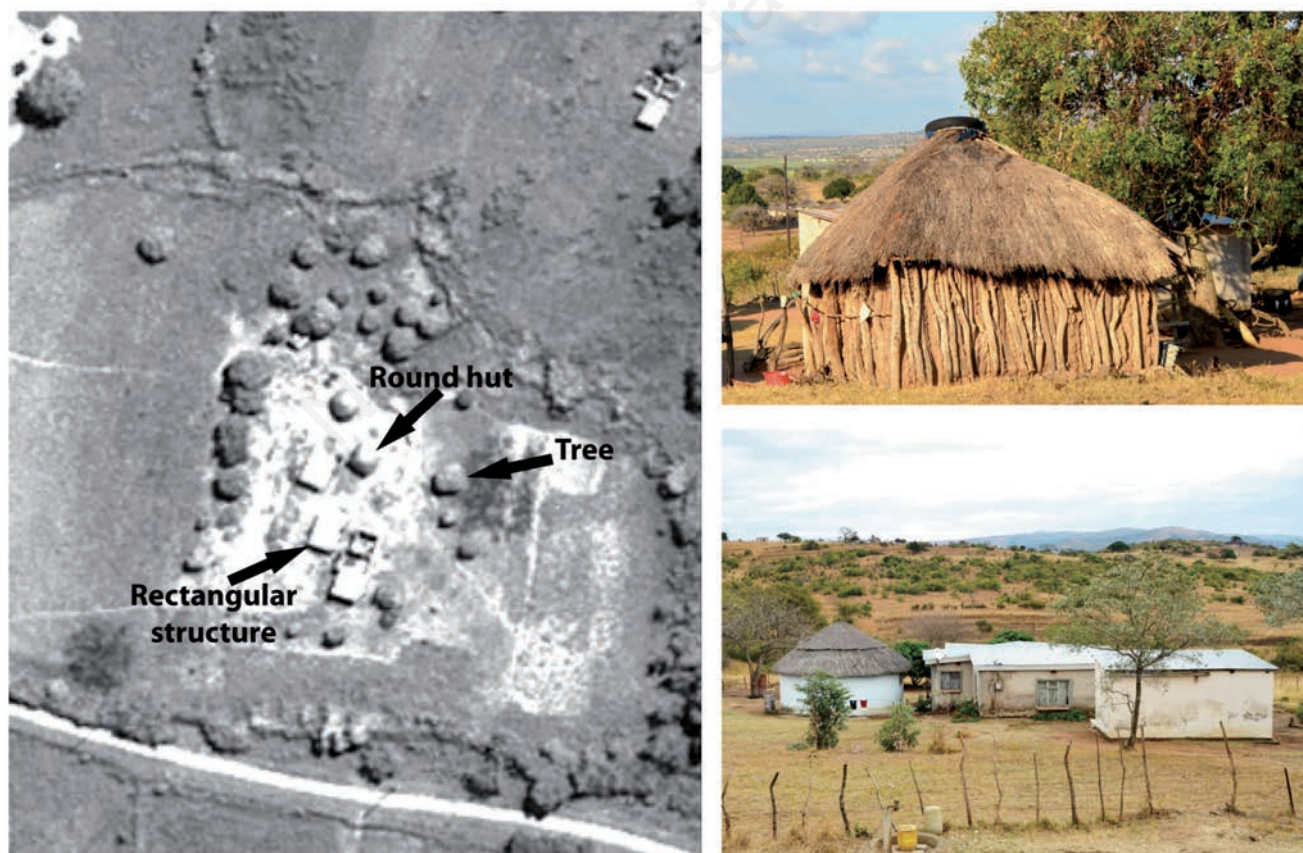

Figure 4. Subset of a panchromatic GeoEye image showing a homestead with round and rectangular buildings including examples of two building structures typical for the study area. ©Geo-Eye, Inc. (2010), provided by E-GEOS S.p.A., under Global Monitoring for Environment and Security Space Component Data Access. Photos: J. Franke.

structures in the three demonstration sites. As a no-cost alternative, Bing Maps was also tested for its use in house mapping. All houses in the VHR sites were manually digitized in ArcGIS (Redlands, CA, USA), and the spatial coverage of VHR data from Bing Maps in these three areas was investigated. At the time, only VHR data for the Mamfene site in South Africa was integrated in Bing Maps. No VHR data were available for the other sites.

Land cover and wetland mapping using high resolution data

Previous studies already demonstrated that certain land cover types correspond with the occurrence of *Anopheles* mosquitos (Pope *et al.*, 1994; Zeilhofer *et al.*, 2007; Dambach *et al.*, 2009). The land cover in the MALAREO study area was assessed by HR RapidEye imagery, since this geo-information can be used directly by the NMCPs for vector control planning and as input for various studies on malaria incidence, vector habitat distribution and population density estimates. The major benefit of using data with 5 m spatial resolution is evident, given how malaria risk often varies within a short distance. For instance, even small water bodies play an important role as larval breeding sites for malaria vectors. RapidEye provides the high level of detail of the final land cover map, which benefits malaria incidence modelling and epidemiological studies in general. This was useful because the area had never been mapped at such high spatial detail before. Small water bodies are a direct indicator for malaria risk and the distance to water is a major factor when modelling malaria incidence. The higher the resolution of the data, the higher the number of small water bodies that can be detected. Since larval breeding is influenced by certain water body characteristics, temporally flooded wetlands, permanent standing and flowing water bodies were differentiated using HR data.

As a first step, a classification scheme is required that includes classes relevant to achieve the project aims. In total, 11 classes were considered of relevance to characterize the project area according to the conditions for the malaria vector and the pattern of human activities (Table 1). As a result, the classification followed a hybrid classification scheme between land cover and land use. In general, class definitions should be as simple as possible, in order to ensure the highest classification accuracies (class complexity vs. class accuracy) and to allow for rapid data processing.

The RapidEye data, mainly acquired in October and November 2011, were first preprocessed using a standardized atmospheric correction in ATCOR (Richter, 1997). The land cover classification of the HR data
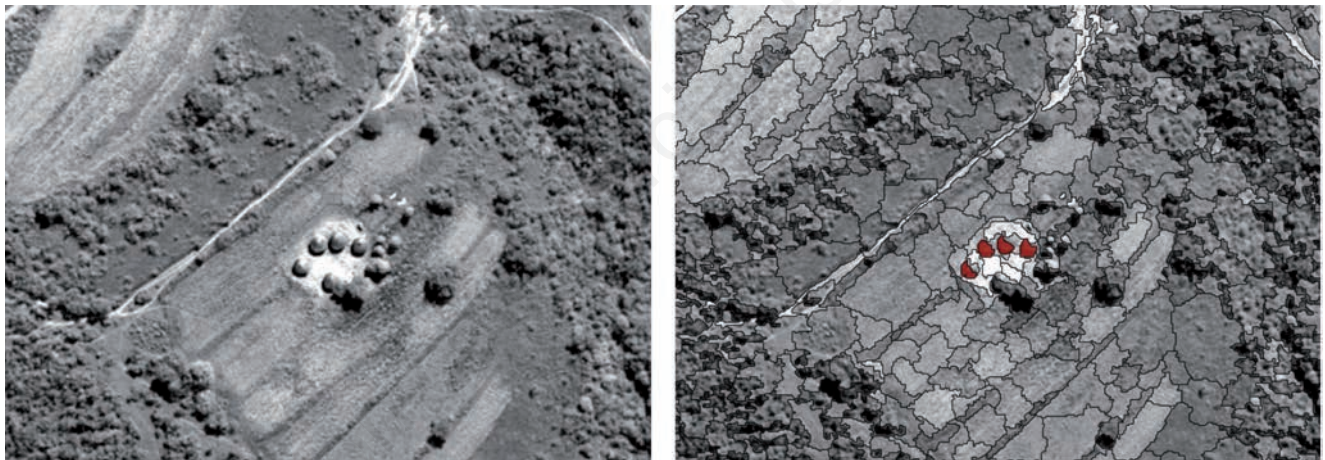

Figure 5. Panchromatic GeoEye image of a rural homestead (left) and the classified image segments with identified huts (right).

Table 1. List of land cover/land uses and their characteristics considered in MALAREO.

Land cover/land use	Characteristics and relevance
Forest/woodland	Tree-dominated land cover
Bush/shrub land	Open/closed bush/shrub-dominated land cover
Grassland/savannah	Open grassland-dominated savannah ecosystem
Standing water	Basins, ponds, lakes *etc.* and water covered wetlands; larval breeding is favoured by non-turbulent water
Flowing water	Mainly rivers and ocean; larval breeding is not favoured by turbulent water
Wetland	Non-permanent, temporally flooded wetland
Large-scale agriculture	Large-scale irrigated sugarcane plantations were recently established in the study area and might favour vector breeding through irrigation
Subsistence farming	Farming plots for subsistence agriculture and cattle breeding
Settlement/infrastructure	Location of population for potential malaria transmission
Bare soil/rock	Mainly mining, rocky outcrops and open sandy areas
Roads/tracks	Transportation infrastructure

was conducted applying an object-based image analysis with a predefined hierarchical rule-set in eCognition software. Object-based classification is often applied for HR/VHR imagery. For this type of data, a pixel-based approach causes a *grainy* appearance of classes over the scene due to single, isolated, class pixels. This effect can be significantly reduced by using object-based approaches.

Since some LULC classes, defined in MALAREO, showed a high class complexity, manual classification steps were necessary (*e.g. large-scale agriculture*, *subsistence farming*, *settlement/infrastructure* and *wetland*). For instance, a temporally flooded wetland can be covered by water, by dense grass or have no vegetation cover at all, which results in wide differences in spectral signatures. These differences within a single land cover class cannot be accounted for by an automated rule-set and require logic that can only be exercised via manual assignment of an object. Although, by strict definition, wetlands include water bodies, this paper defines wetlands as only temporally-flooded wetlands. Water bodies are an independent, separate class.

Due to the fact that land cover was assessed using RapidEye data and image tiles from 5 different image acquisition dates (and that each of the 62 slightly overlapping image tiles were classified separately), differences between class objects could occur in the overlapping areas (*e.g.* due to differences in image segments between tiles). In order to create a seamless classification of the study area, classifications of the image tiles were ultimately fitted by applying a decision fusion approach to each sub-object within overlapping areas showing different class memberships. As a last step, an object-based accuracy assessment of the classification was completed, whereby 270 random image objects were visually interpreted by a project partner uninvolved in the primary classification procedure. The resulting reference land cover classes were compared to the image classification and accuracy metrics were calculated.

Distance to inland water and wetlands

Distance to water is a major determinant of malaria risk. On a local scale, vector densities around reservoirs depend on the distance from the water body and the suitability of the water body to act as a vector breeding habitat (Zeilhofer *et al*., 2007). Besides the fact that anopheline mosquitoes can be displaced over long distances through various factors such as air travel (Tatem *et al*., 2006), the natural flight range of anopheline mosquitoes is estimated at between 0.5 and 2.5 km from their breeding habitats (Russell and Santiago, 1934; Charlwood and Alecrim, 1989; Keiser *et al*., 2005; Zeilhofer *et al*., 2007). As an input for modelling malaria incidence and mapping habitat foci, distance to inland water and distance to wetlands were assessed by calculating the Euclidean distance to the water and wetland layers derived from the HR land cover classification. The Euclidean distance is simply defined as the shortest straight line distance between two points.

Population density estimates

When modelling the spatial distribution and spread of malaria, accurate and detailed information on population distribution and population density is of significant importance. The AfriPop project, launched in July 2009, was initiated with the aim of producing a detailed and freely-available population distribution map for the whole of Africa (http://www.afripop.org). AfriPop has therefore constructed a GIS-linked database with census and official population data. Population count data were combined with detailed, satellite-derived settlement extents to map population distributions across Africa at a finer spatial resolution than ever before (Linard *et al*., 2012). These settlement extents were used to refine existing land cover data sets such as GlobCover (ESA) (http://due.esrin.esa.int/page_globcover.php), provid-

ing significantly improved gridded population datasets across large areas through redistribution of population figures from census and using land cover-based weightings (Tatem *et al*., 2007; Linard *et al*., 2010). In MALAREO, a spatially enhanced and updated AfriPop dataset was generated for the study area (Deleu *et al*., 2015). The HR land cover/use classification allowed a detailed identification of human settlements, particularly of small homesteads in the rural area. The MALAREO settlement extents were used to generate an enhanced and updated version of the AfriPop dataset for the project area with improved spatial detail. The enhanced AfriPop dataset is based on the HR land cover map, census data of 2007 and UN urban/rural growth rates for 2011.

Habitat foci map

Transmission of malaria takes place in the spatial intersection of the vector habitat and the human habitat. Geo-information on these habitat foci allows identifying priority areas for malaria control interventions such as integrated vector management. The population density in proximity to potential vector breeding sites (inland water bodies and wetlands) was therefore analysed by considering the enhanced MALAREO AfriPop data in relation to distance to inland water bodies and wetlands, as given by the HR land cover map. The flight range of anopheline mosquitoes has been estimated at between 0.5 and 2.5 km from their breeding habitats (Russell and Santiago, 1934; Charlwood and Alecrim, 1989; Keiser *et al*., 2005; Zeilhofer *et al*., 2007). The maximum distance displayed in a habitat foci map is adjustable. As a result, this study created multiple habitat foci maps with maximum distance ranging from 0.5 to 2.5 km.

Malaria incidence modelling and vector breeding site analysis

The products described in this paper supported further epidemiological and entomological studies. These included Bayesian modelling for malaria incidence prediction using malaria case data, as well as modelling of potential vector breeding sites and the identification of environmental factors that influence the occurrence of larvae (Dlamini *et al*., 2015).

Results

A number of maps were produced by the study. The final VHR and HR geo-information products generated in MALAREO can all be freely accessed upon request for further use.

Very high resolution household maps

Figure 6 shows an example of the results of the developed house detection approach for the Mamfene test site (ZA). The map shows the distribution of building types in the area with statistical information on type of structure per km^2. These data were made available as GIS layers for the NMCPs. Representatives of 6 southern African countries were trained by MALAREO to use these data in a GIS environment (QGIS), to improve the planning of bed nets distribution and IRS (full area coverage and estimation of required insecticides), and monitoring of their campaign progress. These data were generated for all three VHR sites. Due to the lack of independent reference data on structure type and distribution for a quantitative accuracy assessment, only a qualitative assessment of the accuracy was possible. Due to the integration of land cover information in the classification process, a high level of spatial accuracy and precision of the structure type classification could be achieved and only few false positives were found outside of settle-

ments. A slight over-detection of structures inside of homesteads was observed, since some other infrastructure features were identified as buildings. In terms of thematic accuracy, a non-systematic confusion between round and rectangular houses was partly observed, which suggests general validity of the statistical figures on structure types per area. Table 2 lists the spatial statistics on structure types per area for the VHR test sites as a basis for insecticide consumption estimation and number of required bed nets.

In assessing the cost-free data alternative, spatial coverage of VHR data in Bing Maps was only sufficient in the Mamfene study area. As a result, an additional site at the boarder triangle between Mozambique, Swaziland and South Africa (Naamacha) was mapped in order to assess labour intensity of this manual mapping approach. Depending on the site, digitizing structures using BingMaps in ArcGIS took about 6.2 and 10.3 min per km^2 for an experienced image interpreter. These no-cost/low-cost data sources have some benefits and limitations. Whereas their major benefit is the cost-effectiveness, their limitation is low spatial coverage and partial lack of metadata, like the date of acquisition. However, these free data sources are constantly being updated to improve coverage and are already an alternative image source for some EO applications. However, the current VHR image coverage of these free image sources only sparsely cover rural areas of Africa, which means that these data sources are not yet suitable to support malaria control. Another limitation of the use of such no-cost/low-
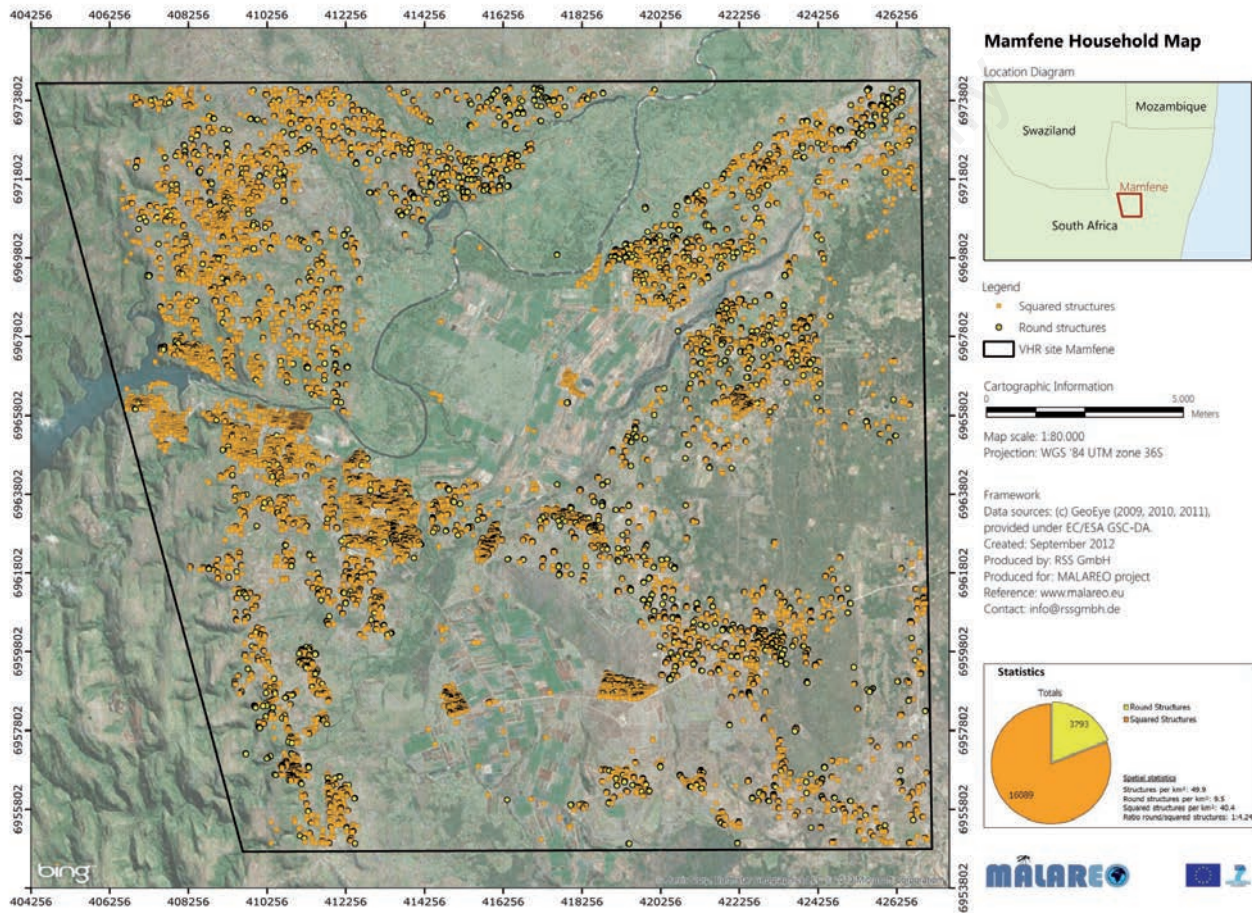

Figure 6. Household map for Mamfene, South Africa based on classified GeoEye data with spatial statistics on the distribution of building structure types. Background: BingMaps.

Table 2. Statistics on building type distribution in the very high resolution test sites.

VHR test site	No. of round structures	No. of squared structures	Buildings per km^2	Round structures per km^2	Square structures per km^2
Mamfene (Jozini Municipality, KwaZulu Natal), South Africa	3793	16,089	49.9	9.5	40.4
Ponta do Ouro (Matutuine district, Maputo Province), Mozambique	58	2826	17.5	0.4	17.2
Dwashini (Hhohho District), Swaziland	4476	11,404	41.4	11.7	29.7
Naamacha, boarder triangle MZ-SZ-ZA	1239	2086	133.0	49.6	83.4

VHR, very high resolution. In the first three test sites VHR data from GeoEye and IKONOS was used; the results from Naamacha are based on BingMaps.

cost data sources is that only manual digitizing of relevant land cover features is possible, which increases personnel costs.

High resolution land cover and wetland maps

Figure 7 shows a subset of the 2011 LULC map based on HR RapidEye data. The map's spatial detail allowed for detection of small-scale LULC features which could substantially improve epidemiological modelling and malaria control management. The accuracy of the land cover map was assessed by a stratified random sampling approach. The results of the accuracy assessment showed an overall accuracy of 80.7% for the land cover classification, with a kappa index of 0.78. The malaria-relevant classes *flowing water* and *large-scale agriculture* demonstrate the highest accuracies (100%). Average user's accuracy was 85.2%. The grassland/savannah class was the least accurate class, with 68.1% accuracy. This is expected to be mainly due to confusion with the class bush-/shrub land, and by the continuous transition of these land covers in the area. This class confusion substantially decreased the overall accuracy of the land cover classification, since most of the other classes showed higher accuracies than 80%. Table 3 shows the land cover statistics in the study area, which is dominated by forest/woodland, bush-/shrub land and grassland/savannah.

The wetland map of the whole study area is displayed in Figure 8. Due to the high spatial resolution of the RapidEye data, even small wetlands could be identified that could act as vector breeding sites. In total, 3046 standing water bodies and 10,503 non-permanently flooded wetlands were identified. This layer was used as input for the distance to wetland calculation, modelling of malaria incidence (Giardina *et al.*, 2015), habitat foci mapping and assessment of potential vector breeding sites (Dlamini *et al.*, 2015).

Distance to inland water and wetlands

The Euclidian distance to inland water (standing and flowing water) and the distance to wetlands were calculated for the study area as input for the modelling of malaria incidence, since distance to potential breeding sites is a major determinant. Figure 9 shows the distance to inland water map (in meters).

Enhanced AfriPop population density map

The AfriPop approach was applied in MALAREO by the use of the HR land cover map and census data of 2007, whereby the UN urban/rural growth rates for 2011 were additionally considered in order to account for recent population trends (Deleu *et al.*, 2015). The result was a spatially enhanced and updated AfriPop dataset for the study area.

Habitat foci map

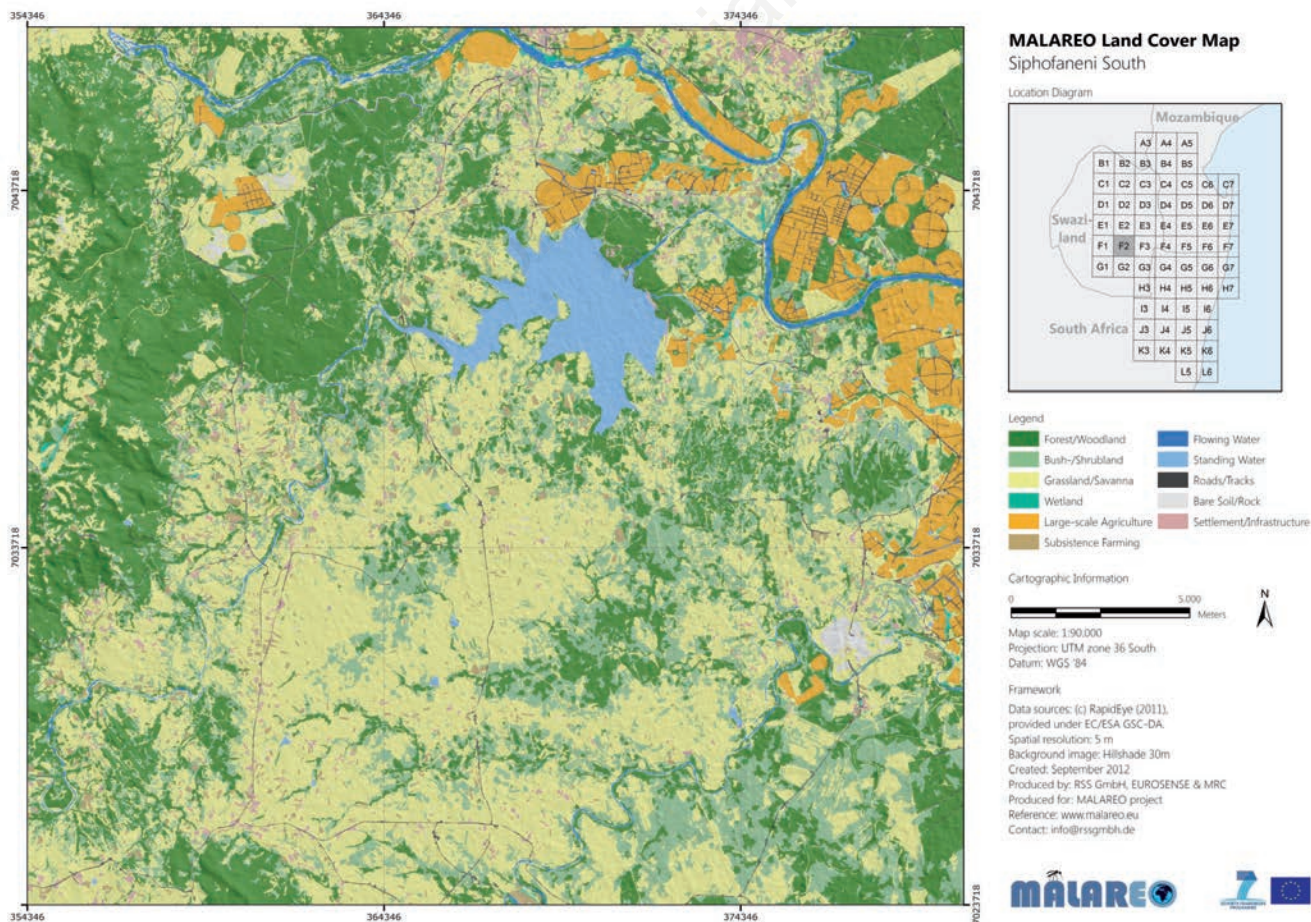Population density within the spatial intersection of potential vector

Figure 7. Subset of the high resolution land cover/land use map for an area in Swaziland.

Table 3. Land cover statistics of the study area.

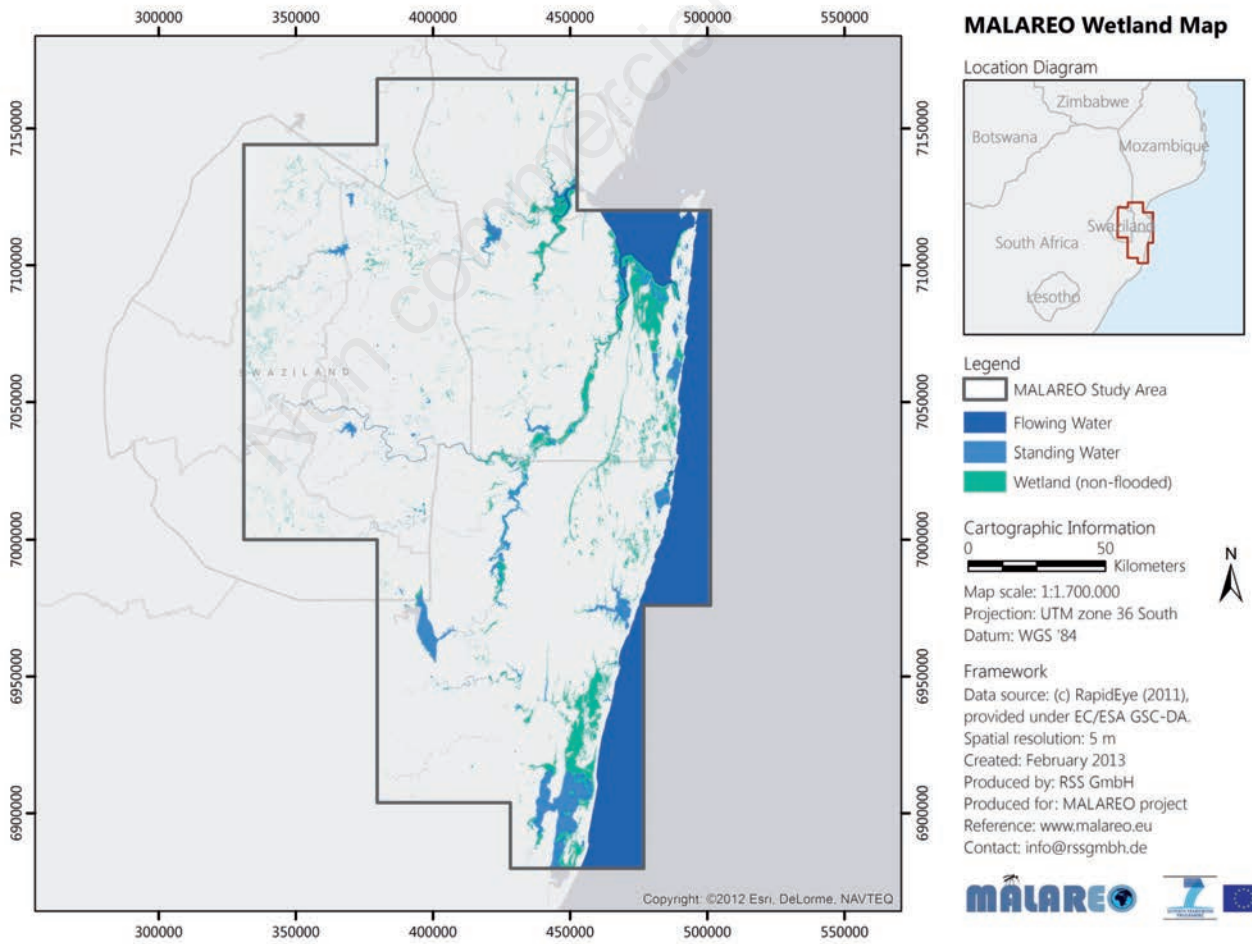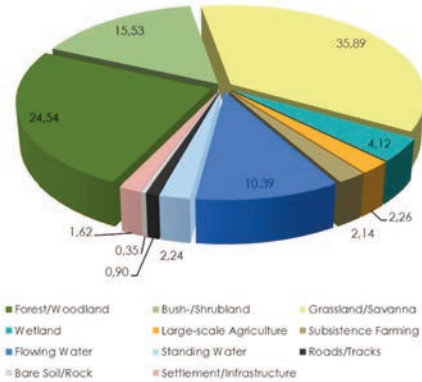| LC class	Area (km^2)	Area (%)
Forest/woodland	8617.8	
Bush-/shrub land	5454.5	
Grassland/savannah	12,602.5	
Wetland	1448.2	
Large-scale agriculture	795.0	
Subsistence farming	751.3	
Flowing water	3648.9	
Standing water	784.8	
Roads/tracks	316.0	
Bare soil/rock	122.7	
Settlement/infrastructure	569.8	
Sum	35,111.5	

LC, land cover.

Figure 8. High resolution wetland map that differentiates flowing and standing water bodies as well as non-flooded wetlands.

habitat and human habitat was investigated using the inland water bodies and wetland layer, and the enhanced, updated AfriPop data. The area of intersection was determined by applying various distances from wetlands within the flight range of the anopheline mosquitoes, resulting in different habitat foci maps. Figure 10 shows the population density within a 1km buffer to major inland water bodies as identified with RapidEye imagery. This analysis showed that about 1.78 million people live in the study area, of which 28% (0.47 million) live within a 1 km distance to major inland water bodies. The habitat foci map indicates priority areas for malaria control interventions.

Discussion

The key technology required to support the malaria vector control and management programmes, as well as malaria epidemiology through EO products, is the range of existing satellites. In the past, EO sensor abilities limited the use of EO data in epidemiology and disease control. However, today there are numerous sensors with improved spatial, temporal and spectral characteristics, and multiple missions are planned to launch additional, high-performance sensors into space.

The combination of remote sensing and GIS has strengthened the opportunities for disease surveillance, control and early warning. The results demonstrate the potential of HR geo-information for the improvement and support of malaria control management and epidemiological studies. Although the ability to enhance malaria control through GIS-based management systems has already been demonstrated (Booman et al., 2003), this project added additional HR and VHR EO components.

Low performance of the EO sensors available in the past, high data costs and long product turnaround times, as well as previous remote sensing data or statistical models with low spatial resolution, meant studies were only carried out in small areas with impractical time frames for updates. The last few decades have shown a substantial increase in potential EO approaches and knowledge regarding environmental parameters that influence spatial and temporal disease risk patterns. In 1998, Hay et al. provided an excellent review of earlier available remote sensing techniques for malaria epidemiology that can now be compared with today's technologies. The current and future range of satellite systems enables the development of new and innovative EO services for health applications (Rogers et al., 2002; Ceccato et al., 2005). As demonstrated in MALAREO with HR and VHR data, the spatial detail of disease risk and their environmental factors could be
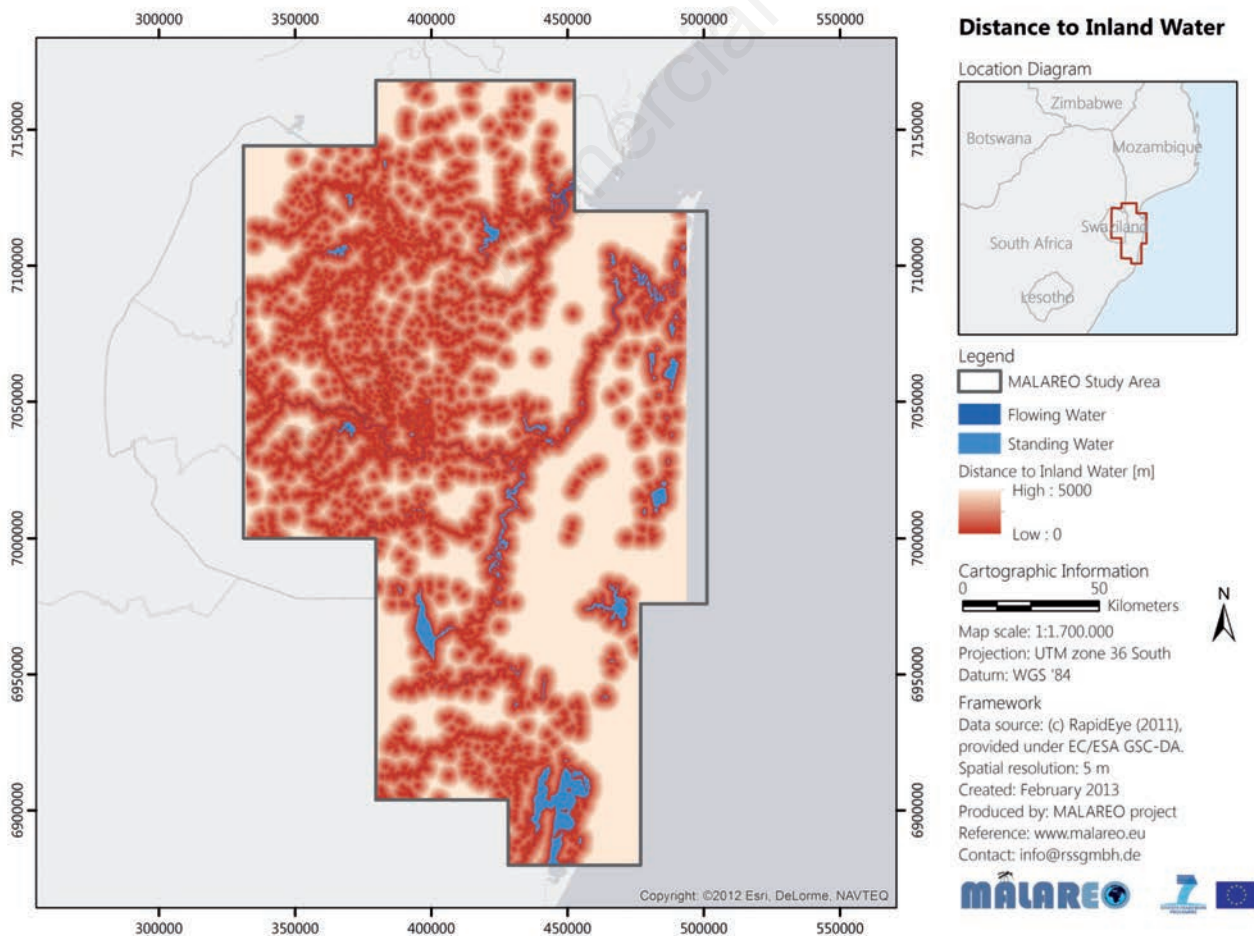

Figure 9. Distance to inland water map.

improved substantially, which can vary greatly across small distances. With the use of HR data at a national scale and with frequently scheduled monitoring, such geo-information products can directly support the health sector in Africa.

Users from the NMCPs in the study area were involved in the development of these products by defining their requirements. All final products were provided for integration in their working environment and management systems. Besides the development and provision of these EO products, a major issue is the capacity development in the NMCPs that is required to ensure that this geo-information can be used effectively. Therefore, apart from teaching local staff the processes behind EO product development, a major requirement is to provide training in the use of GIS, geostatistics and how to interpret EO products. In MALAREO, three workshops were held for end users of the NMCPs of southern African countries with the aim of transferring knowledge, whereby feasibility in terms of data, methodologies, financial resources, hard- and software and improvement of the existing systems/methods (using open-sources software) were taken into account. In general, the overall future approach should combine user requirements, existing capacity and funding strategies to develop long-term perspectives.

EO data costs are another factor in determining the sustainability of

remotely sensed data in malaria control. Low to medium resolution data are available free of charge and can already be used for some disease-related applications. For instance, data from satellite sensors such as MODIS can be used for climatic suitability assessments and medium resolution data from Landsat-8 (30 m) (*e.g.* via http://www.usgs.gov/) can be used for land cover mapping. However, particularly in the context of malaria, there is a major benefit of HR and VHR data, since small land cover features (*e.g.* small water bodies, wetlands, houses and huts in rural areas, *etc.*) often play major roles in the malaria transmission cycle. HR and VHR data are particularly suitable for mapping small-scale variations in determining factors of malaria risk. The potential of VHR images was tested in the present study, even if they are still too expensive for operational use in malaria control. Though promising results could be achieved for parts of the study area with free image services (Google Earth, BingMaps *etc.*), but better coverage is required. It might, however, be a realistic alternative if coverage and updating improve in the future. The RapidEye system proved to be very useful for malaria control and epidemiological studies, but even their comparably low costs of 1.28 USD/km^2 (see http://blackbridge.com/rapideye/index.html for pricing updates) are likely to exceed the budgets of the NMCPs in some countries in southern Africa. However, new HR earth observation missions such as the Sentinels
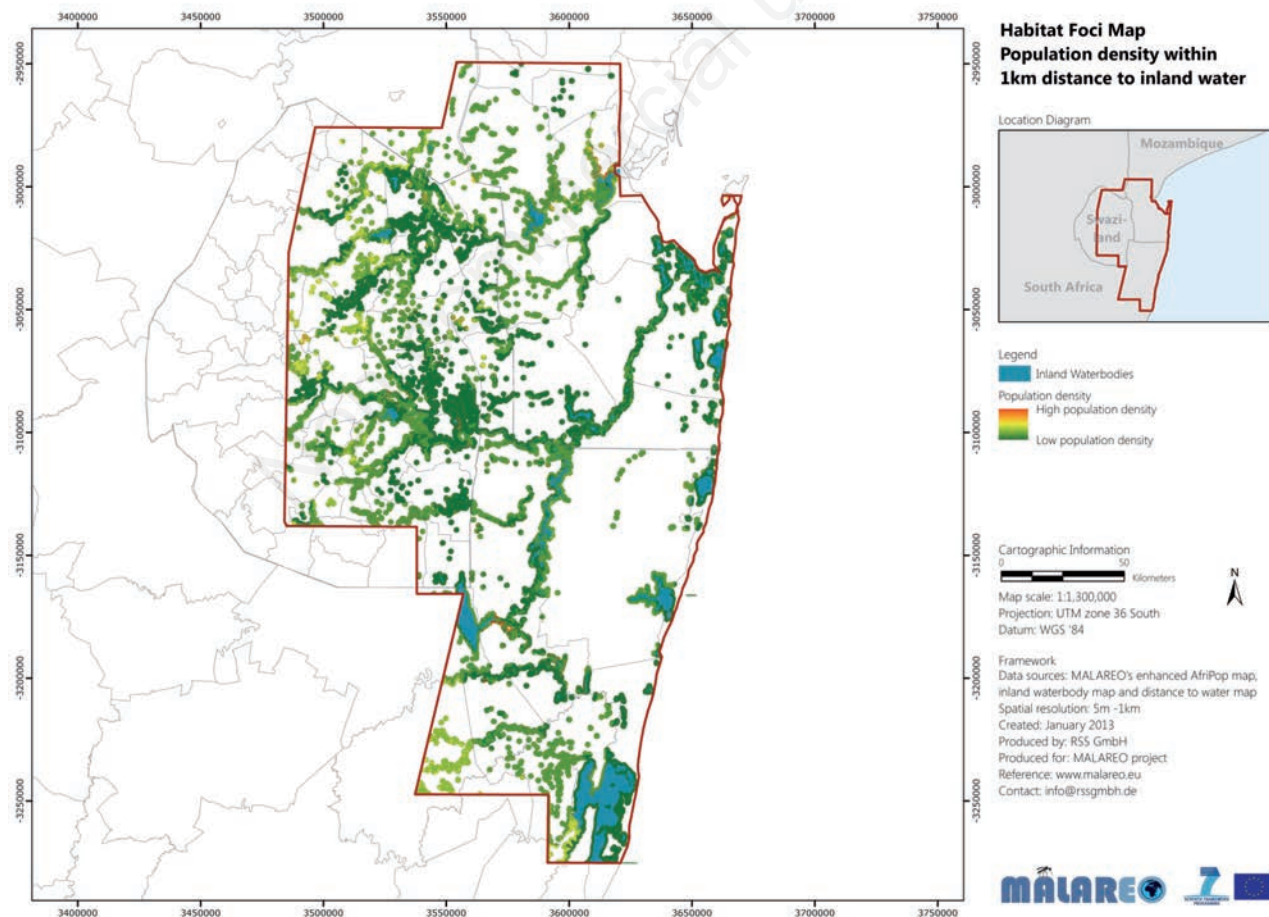

Figure 10. The habitat foci map shows the population density within the distance of the flight range of anopheline mosquitos from potential breeding sites. This example shows the population density within a 1 km buffer to inland water bodies.

(http://www.esa.int/) have a free and open data policy (ESA and the European Commission), which will further encourage the use of EO techniques in the health sector.

Handling and processing vast amounts of HR data requires a high level of automization. Standardized, reliable and timely geo-information must be derived from EO data that can be integrated in spatial decision support systems (SDSS). Such systems should be developed in order to extend EO capacity into operational disease surveillance, control and early warning with a generic focus on resource-poor environments and architectures. Kelly *et al*. (2012) and Shaw (2012) have requested further development, validation and sustainable implementation of SDSS tools in operational disease control. The operationalization of the service generation within a SDSS will be a key for providing up-to-date information on disease risk. This project provided a range of EO-based geo-information products for vector control and management as well as malaria epidemiology, helping to better understand malaria-related environmental factors and directly support malaria control management. Through targeted capacity development, the use of the project results could be sustained beyond the project duration. Monitoring techniques such as those demonstrated in MALAREO need to be disseminated among the funding agencies supporting the NMCPs and among malaria control managers, in order to raise awareness of new opportunities to improve malaria eradication efforts.

Conclusions

The monitoring applications developed in the MALAREO project added a HR EO component to existing malaria control strategies in southern Africa. Earth observation monitoring approaches were designed to meet the user needs of the NMCPs in South Africa, Swaziland and Mozambique, and the technical know-how needed was developed, resulting in the introduction of open-source GIS and the use of the MALAREO EO products. The outcomes of the project have high potential for operationalization, transferability to other countries and replicability for studies on other environmental related diseases. High resolution data from the Sentinel mission will continue to advance the use of EO techniques in malaria control and epidemiology, through its improved technical capabilities as well as free and open data policy. With this emerging suitable satellite technology, further efforts are necessary to raise awareness among decision-makers concerning EO benefits, and ultimately establishing EO as a ubiquitous tool for malaria monitoring.

References

African Union, 2007. Advocacy strategy document presented at the 3rd session of the AU conference of Ministers of Health. Africa Malaria Elimination Campaign ed., Johannesburg, South Africa.

Balfour T, 2002. TB and malaria in SADC countries. In Petrida L, ed. South African health review 2002. Health Systems Trust, Durban, South Africa, pp 305-26. Available from: http://www.hst.org.za/uploads /files/chapter16.pdf

Beck LR, Bradley ML, Wood BL, 2000. Remote sensing and human health: new sensors and new opportunities. Emerg Infect Dis 6:217-66.

Bejon P, Williams TN, Liljander A, Noor AM, Wambua J, Ogada E, Olotu A, Oiser FHA, Hay SI, Färnert A, Marsh K, 2010. Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya. PLoS Med 7:e1000304.

Booman M, Sharp BL, Martin CL, Manjate B, La Grange JJ, Durrheim DN, 2003. Enhancing malaria control using a computerised management system in southern Africa. Malaria J 2:13.

Bousema T, Drakeley C, Gesase S, Hashim R, Magesa S, Mosha F, Otieno S, Carneiro I, Cox J, Msuya E, Kleinschmidt I, Maxwell C, Greenwood B, Riley E, Sauerwein R, Chandramohan D, Gosling R, 2010. Identification of hot spots of malaria transmission for targeted malaria control. J Infect Dis 201:1764-74.

Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, Ghani A, Drakeley C, Gosling R, 2012. Hitting hotspots: spatial targeting of malaria for control and elimination. PLoS Med 9:e1001165.

Ceccato P, Connor PJ, Jeanne I, Thomson MC, 2005. Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. Parasitologia 47:81-96.

Ceccato P, Gobron N, Flasse S, Pinty B, Tarantola S, 2002. Designing a spectral index to estimate vegetation water content from remote sensing data. Part 1: theoretical approach. Remote Sens Environ 82:188-97.

Charlwood J, Alecrim WA, 1989. Capture-recapture studies with the South American malaria vector *Anopheles darlingi*. Root. Ann Trop Med Parasit 83:569-76.

Crombie MK, Gilles RR, Arvidson RE, Brookmeyer P, Weill GJ, Sultan M, Harb M, 1999. An application of remotely derived climatological fields for risk assessment of vector-borne disease: a spatial study of Filariasis prevalence in the Nile delta, Egypt. Photogramm Eng Rem S 65:1401-9.

Curran PJ, Atkinson PM, Foody GM, Milton EJ, 2000. Linking remote sensing, land cover and disease. Adv Parasit 47:37-80.

Dambach P, Sié A, Lacaux JP, Vignolles V, Machault V, Sauerborn R, 2009. Using high spatial resolution remote sensing for risk mapping of malaria occurrence in the Nouna district, Burkina Faso. Global Health Action 2:10.3402/gha.v2i0.2094.

Deleu J, Franke J, Gebreslasie M, Linard C, 2015. Improving AfriPop dataset with settlement extents extracted from RapidEye for the border region South-Africa, Swaziland and Mozambique. Geospatial Health (in press).

Dlamini SN, Franke J, Vounatsou P, 2015. Assessing the relationship between environmental factors and malaria vector breeding sites in Swaziland using multi-scale remotely sensed data. Geospatial Health (in press).

Giardina F, Franke J, Vounatsou P, 2015. Geostatistical modeling of malaria risk in Mozambique: assessing the effect of remotely-sensed imagery spatial resolution. Geospatial Health (in press).

Guerra CA, Gikandi PW, Tatem AJ, Noor AM, Smith DL, Hay SI, Snow RW, 2008. The limits and intensity of plasmodium falciparum transmission: implications for malaria control and elimination worldwide. PLoS Med 5:300-11.

Hay SI, Snow RW, Rogers DJ, 1998. From predicting mosquito habitat to malaria seasons using remotely sensed data: practice, problems and perspectives. Parasitol Today 14:306-13.

Kalluri S, Gilruth P, Rogers D, Szczur M, 2007. Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review. PLoS Pathog 3:1361-71.

Keiser J, Caldas de Castro M, Maltese MF, Bos R, Tanner M, Singer BH, Utzinger J, 2005. Effect of irrigation and large dams on the burden of malaria on a global and regional scale. Am J Trop Med Hyg 72:392-406.

Kelly GC, Tanner M, Vallely A, Clements A, 2012. Malaria elimination:

moving forward with spatial decision support systems. Trends Parasitol 28:297-304.

Kunene S, Phillips AA, Gosling RD, Kandula D, Novoyny JM, 2011. A national policy for malaria elimination in Swaziland: a first for sub-saharan Africa. Malaria J 10:313.

Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ, 2012. Population distribution, settlement patterns and accessibility across Africa in 2010. PLoS One 7:e31743.

Linard C, Gilbert M, Tatem AJ, 2010. Assessing the use of global land cover data for guiding large area population distribution modelling. Geoj Lib 76:525-38.

MISAU, 2011. Moçambique inquérito demográfico e de saúde 2011. Ministério da Saúde, Maputo, Mozambique.

Moonasar D, Nuthulaganti T, Kruger PS, Mabuza A, Rasiswi ES, Benson FG, Maharaj R, 2012. Malaria control in South Africa 2000-2010: beyond MDG6. Malaria J 11:294.

Pope KO, Rejmankova E, Savage HM, Arredondo-Jimenez JI, Rodriguez MH, Roberts DR, 1994. Remote-sensing of tropical wetlands for malaria control in Chiapas, Mexico. Ecol Appl 4:81-90.

Richter R, 1997. Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. Int J Remote Sens 18:1099-11.

Rogers DJ, Randolph SE, Snow RW, Hay SI, 2002. Satellite imagery in the study and forecast of malaria. Nature 415:710-5.

Russell PF, Santiago D, 1934. Flight range of the *Funestus-Minimus* subgroup of *Anopheles* in the Philippines. Am J Trop Med 14:139-57.

Sharp BL, Kleinschmidt I, Streat E, Maharaj R, Barnes KI, Durrheim DN, Ridel FC, Morris N, Seocharan I, Kunene S, La Grange JP, Mthembu JD, Maartens F, Martin CL, Barreto A, 2007. Seven years of regional malaria control collaboration. Mozambique, South Africa, and Swaziland. Am J Trop Med Hyg 76:42-7.

Shaw NT, 2012. Geographical information systems and health: current state and future directions. Healthc Inform Res 18:88-96.

Tatem AJ, Noor AM, von Hagen C, Di Gregorio A, Hay SI, 2007. High resolution population maps for low income nations: combining land cover and census in East Africa. PLoS One 2:e1298.

Tatem AJ, Noor AM, Hay SI, 2004. Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery. Remote Sens Environ 93:42-52.

Tatem AJ, Rogers DJ, Hay SI, 2006. Estimating the malaria risk of African mosquito movement by air travel. Malaria J 5:57.

Taubenböck H, Esch T, Wurm M, Roth A, Dech S, 2010. Object-based feature extraction using high spatial resolution satellite data of urban areas. J Spat Sci 55:117-33.

Tyc G, Tulip J, Schulten D, Krischke M, Oxfort M, 2005. The RapidEye mission design. Acta Astronaut 56:213-9.

WHO, 2012a. WHO guide on international travel and health. World Health Organization, Geneva, Switzerland. Available from: http://www.who.int/ith/en/.

WHO, 2012b. World malaria report: 2012. World Health Organization, Geneva, Switzerland. Available from: http://www.who.int/malaria/publications/world_malaria_report_2012/report/en/index.html.

WHO, 2013a. Vector control of malaria. World Health Organization, Geneva, Switzerland. Available from: http://www.who.int/malaria/areas/vector_control/en/.

WHO, 2013b. Vectors, environment and society research. World Health Organization, Geneva, Switzerland. Available from: http://www.who.int/tdr/research/vectors/en/.

Zeilhofer P, dos Santos ES, Ribeiro ALM, Miyazaki RD, dos Santos MA, 2007. Habitat suitability mapping of *Anopheles darlingi* in the surroundings of the Manso hydropower plant reservoir, MatoGrosso, Central Brazil. Int J Health Geogr 6:7.