

A climate distribution model of malaria transmission in Sudan

Mohammed I. Musa^{1,2}, Shamarina Shohaimi¹, Nor R. Hashim³, Isthrinayagy Krishnarajah^{4,5}

¹Department of Biology, Faculty of Science, Putra University, 43400 Serdang, Selangor, Malaysia; ²Economic and Social Research Bureau, Ministry of Science and Technology, Khartoum, Sudan; ³Department of Environmental Sciences, Faculty of Environmental Studies, Putra University, 43400 Serdang, Selangor, Malaysia; ⁴Department of Mathematics, Faculty of Science, Putra University, Serdang, 43400 Selangor, Malaysia; ⁵Laboratory of Computational Statistics and Operations Research, Institute for Mathematical Research, Putra University, Serdang, 43400 Selangor, Malaysia

Abstract. Malaria remains a major health problem in Sudan. With a population exceeding 39 million, there are around 7.5 million cases and 35,000 deaths every year. The predicted distribution of malaria derived from climate factors such as maximum and minimum temperatures, rainfall and relative humidity was compared with the actual number of malaria cases in Sudan for the period 2004 to 2010. The predictive calculations were done by fuzzy logic suitability (FLS) applied to the numerical distribution of malaria transmission based on the life cycle characteristics of the *Anopheles* mosquito accounting for the impact of climate factors on malaria transmission. This information is visualized as a series of maps (presented in video format) using a geographical information systems (GIS) approach. The climate factors were found to be suitable for malaria transmission in the period of May to October, whereas the actual case rates of malaria were high from June to November indicating a positive correlation. While comparisons between the prediction model for June and the case rate model for July did not show a high degree of association (18%), the results later in the year were better, reaching the highest level (55%) for October prediction and November case rate.

Keywords: fuzzy logic, climate suitability, malaria rate, geographical information system, mapping, Sudan.

Introduction

Malaria is a serious global health problem, which claims thousands of lives each year. The disease is generally limited by environmental and climatic factors with temperature, rainfall and relative humidity considered to be the most important factors influencing its transmission and large-scale distribution (Cox et al., 1999; Hardy, 2003).

Temperature is critical for the transmission of malaria through its effect on sporogonic duration and mosquito survival (Onori and Grab, 1980). Parasite growth ceases at 16 °C and transmission below 18 °C is unlikely because few adult mosquitoes survive the 56 days needed for sporogony at that temperature (Adjuik et al., 1998). However, at 22 °C sporogony can be concluded in less than three weeks and mosquito survival level is then good enough for the transmission cycle to be completed. The upper temperature for transmission is determined by the vector as temperatures above 32 °C lead to high

turnover of vector populations resulting in weak individuals and high mortality (Craig, 2009). Thermal-induced mosquito death occurs between 40 °C and 42 °C depending on the species (Jepson et al., 1947). Importantly, the higher the temperature is below this level, the more rapid the mosquito developmental cycle resulting in reduced intervals between blood meals and increased frequency of host-vector contact (Gillies and De Meillon, 1968).

Although rainfall does not affect the parasite directly, it plays a crucial role for malaria transmission by providing the medium for the aquatic mosquito stages. It also increases relative humidity, which is important for the activity and survival of all anopheline mosquitoes. Indeed, an average monthly relative humidity below 60% shortens the life of the mosquito enough to make it unsuitable for malaria transmission (Pampana, 1969). The optimal amount of rainfall for transmission has been discussed at length in the project on "Mapping Malaria Risk in Africa" (MARA) indicating that the requirement for steady transmission is around 80 mm in northern Africa, where high temperatures result in a particularly rapid growth rate of mosquito populations (Adjuik et al., 1998; Craig et al., 1999). In this area three months of rain is sufficient for transmission, while the required period is five months in the rest of

Corresponding author:
Shamarina Shohaimi
Faculty of Science, Putra University
43400 Serdang, Selangor, Malaysia
Tel. +60 3 8946 6637; Fax +60 3 8656 7454
E-mail: shamarina@putra.upm.edu.my

the continent. Excessive rainfall, however, is negative for the transmission cycle as it can flush out the mosquito larvae (Martens et al., 1995). Thus, temperature, rainfall as well as relative humidity must be kept at satisfactory levels and their fluctuation on a monthly basis needs to be periodic to sustain transmission.

The aim of this study was to create “suitability” maps by applying fuzzy logic (Zadeh, 1965), which is an extension of Boolean logic dealing with datasets that do not result in a clear outcome but rather produce fuzzy, continuous datasets ranging from one extreme to the other. The fuzzy logic suitability (FLS) model is based on degrees of truth rather than simply true (1) or false (0) just as there are no clear boundaries separating malarious from non-malarious regions (McMichael, 2003). Thus, FLS-classified maps do not only indicate where the climate is suitable vis-à-vis unsuitable, but also indicate continuous changes of chance between these endpoints (Craig et al., 1999).

Materials and methods

To facilitate the description of the procedure followed, the various steps described below are displayed in a flowchart (Fig. 1).

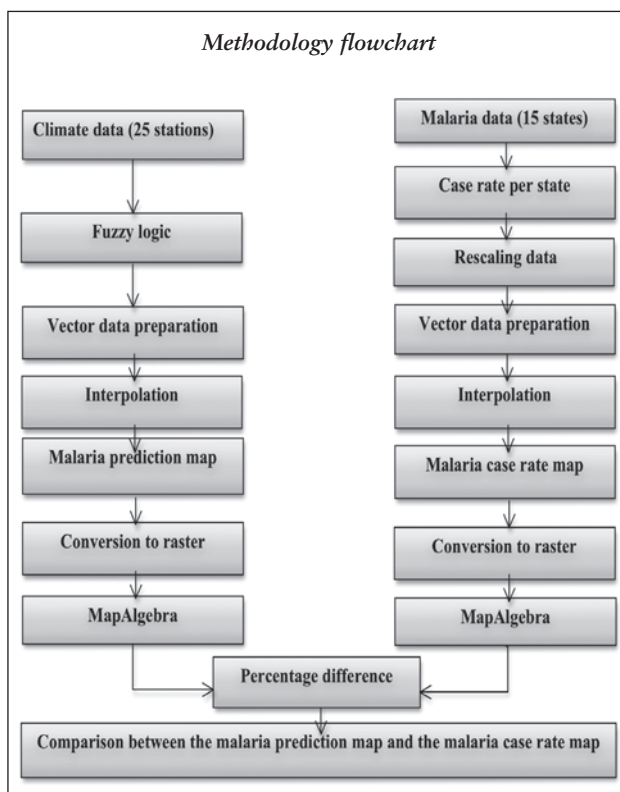


Fig. 1. Flowchart describing the work step by step.

Study area and general climate

The centre of Sudan is geographically located in North-East Africa around latitude 15° N and longitude 30° E. Fig. 2 shows the state borders and the distribution of meteorology stations. The climate ranges from arid in the North to tropical dry in the South with few permanent water bodies apart from the River Nile. Sudan is in dry season from January to March due to north-easterly winds from the Arabian Peninsula. There is no rainfall during this period except in a small area in north-eastern Sudan, where the winds from the Mediterranean bring light rains. In early April, the weather changes for moist, south-western winds bringing heavy rains and thunderstorms to southern Sudan (Metz, 1992). By July, this pattern has covered most of southern and central Sudan to reach the North in August. In some years, humid air may also reach the Egyptian border. In September, the dry north-easterly winds strengthen and push southwards to cover the entire country by the end of December. The capital of Sudan, Khartoum, has a 3-month rainy season (July to September) with an annual average rainfall of 160 mm. Temperatures peak at the end of the dry season. During the warmest months (May and June), the temperature can reach 48 °C contributing to average highs of 41°C. Northern Sudan, with its short rainy season, has the highest average temperature apart from the winter months in the North-East, where there is rain in January and February (Metz, 1992).

Data collection

Climate data

Climate data for the period of 2004-2010, including average monthly maximum and minimum temperatures, rainfall and relative humidity, recorded by the Sudan Meteorological Authority (SMA) were collected from 25 different weather stations. These sites, with their geographical coordinates and elevation specified, are randomly distributed throughout Sudan with a concentration in the centre of the study area (Fig. 2).

Malaria data

Malaria data, assembled by the national malaria control programme (NMCP) of the Federal Ministry of Health (FMH), Sudan, were collected for the period of 2004-2010 from health centres and hospitals and aggregated for each of the 15 states of the country. A map with the Lambert conformal conic projec-

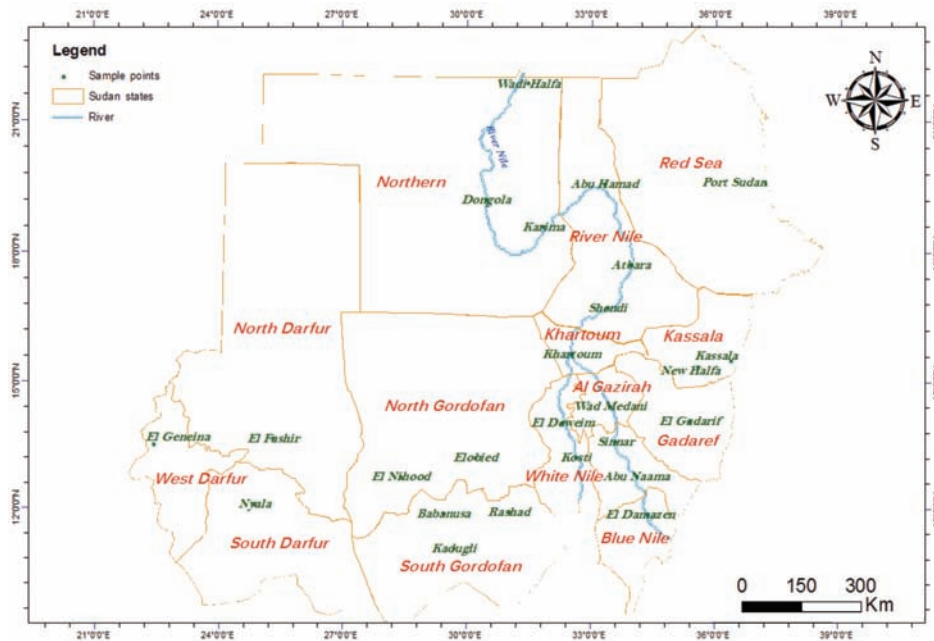


Fig. 2. State borders and meteorology stations in Sudan.

tion, scale 1:5,650,000 (<http://www.mappery.com>, 2009) was used as base for the maps produced.

The fuzzy logic suitability approach

The study adopted the approach of numerical distribution of malaria transmission based on life cycle characteristics of the *Anopheles* mosquito. The FLS model was used for prediction of the impact of climate factors (maximum and minimum temperatures, rainfall and relative humidity) on malaria transmission. The fuzzy logic model was applied as there are no specific boundaries separating malarious regions from non-malarious regions. Likewise, the climatic factors follow spatial gradients, so the occurrence of transmission cannot be defined by cut-offs.

Instead of conditions such as true (*T*) or false (*F*), the fuzzy function were used to transfer the climate data to “climate suitability maps” utilizing fractions between zero (unsuitable conditions = *F*) and 1 (suitable conditions = *T*). The simple, sigmoid, fuzzy membership curve (Eq. 1), adopted from Cox et al. (1999), is defined in MATLAB (<http://www.scribd.com/doc/39964212/Fuzzy-Logic-MATLAB>) as:

$$Y = \text{COS}^2 \left[\frac{(X-F)}{(T-F)} * (\text{Pi}/2) \right] \quad (\text{Eq. 1})$$

where *Y* is the fuzzy membership value of the climate value, *T* the value of *X* when conditions are completely suitable for malaria transmission and *F* the value of *X* when conditions are completely unsuitable. For

rainfall: *F* = 0 and *T* = 80 mm per month; for humidity: *F* = 0 and *T* = 60% per month; for minimum temperature: *F* = 18 °C and *T* = 22 °C; and for maximum temperature: *T* = 32 °C and *F* = 40 °C.

Sigmoid, fuzzy membership curves were employed to classify the climate variables derived from the spatially interpolated weather station data (Hutchinson et al., 1996). For optimum transmission conditions for malaria, the temperature, rainfall and relative humidity need not only be suitable, but must also occur simultaneously (Craig et al., 1999). The number of months meeting threshold conditions was included as parameter in the model, since the ecological differences and variations in vector behaviour require a much shorter period of optimum conditions (three months) to build up sufficient vector populations in the more arid regions in northern Africa than elsewhere (Omumbo et al., 2004).

The malaria case rate

The number of malaria cases, expressed as a function of population and area, i.e. the malaria rate, was calculated in for each states according to Eq. 2 below. This was done to transfer the malaria rate into the ArcMap 10 software (ESRI, Redlands, CA, USA) and produce the malaria rate per state per month.

$$\text{malaria rate} = \frac{\text{number of malaria cases (state)}}{\text{total population (state)}} \quad (\text{Eq. 2})$$

Rescaling

By adding or subtracting a constant and then multiply or divide by another constant, the units of measurement of the data can be changed. Scaling 0-1 means that each variable in the dataset was recalculated as $(\text{value} - \text{min})/(\text{max} - \text{min})$ with “value” representing the malaria rate, and “min” and “max” the minimum and maximum malaria rates, respectively. This method creates at least one observed value at the endpoints (0 and 1).

Vector importation

All feature classes were in the GCS_WGS_1984 coordinate system. Rivers, cities and states layers were all imported into the ArcMap 10 software. The total area of Sudan was georeferenced by specifying four known coordinate points and adjusting the corresponding software points. Each state was digitized as a polygon in which points corresponded to the metrological stations and lines to the rivers. The Lambert conformal conic projection of the reference map was adopted and geostatistical analysis applied to produce the malaria suitability and case rate maps.

Thin plate spline interpolation

The term thin plate spline (TPS) originally refers to the physical analogy of bending a thin sheet of metal but the technique has since found mathematical applications such as smooth multivariate interpolation of irregularly scattered noisy data (Hutchinson et al., 1996; Hutchinson, 1998). TPS is a deterministic, locally stochastic interpolation, which is useful and robust in areas with irregularly spaced data points (New et al., 2002). The usefulness of TPS for dealing with climate data was first shown by Wahba (1983) and Hutchinson and Bischof (1983). Recently, Hutchinson et al. (2009) used TPS to develop long-term, spatial models for daily minimum and maximum temperatures and precipitation for Canada. The main advantage of splines compared to other geo-statistical methods is that prior estimation of the spatial auto covariance structure is not required (Hutchinson, 1998).

The observations by meteorological stations in Sudan, including the monthly temperature maximum and minimum, rainfall and relative humidity, were interpolated to produce climate data covering the whole country for the study period. The malaria rates were also interpolated using this method. The TPS interpolation techniques adopted for this study gener-

ated a surface passing through the control points (CPs) with the least possible change in slope overall. In other words, TPS fitted the CPs with a minimum curvature of the surface. These TPS techniques have the capability of modelling spatial distributions as a function of observational data across an area without prior knowledge of the distribution, thereby achieving spatial accuracy through a minimum of error interpolation.

Conversion to raster

The entire set of geo-statistical analysis layers (GA), including prediction maps and case rate maps, were converted into raster format with a cell size of 6.07 x 6.07 km. MapAlgebra, an ArcGIS geoprocessing tool, was then used to calculate the differences between the two maps as follows:

$$\text{percentage difference} = [(\text{Map1} - \text{Map2}) / (\text{Map1} + \text{Map2})/2] \times 100 \quad (\text{Eq. 3})$$

Results and discussion

Two different malaria maps were produced, one prediction map and one malaria case rate map. However, there is a degree of ambiguity in the latter variable as it excludes the small, but unknown number, of people without clinical disease. The “suitability climate model” presented was used to define the crude distribution of malaria transmission in Sudan. We feel that the datasets and the methodological approach are appropriate at the scale investigated, even if small-scale anomalies that may affect the distribution (e.g. the Nile flood-plains in areas of low rainfall, the Saharan area, agricultural practices, deforestation, etc.) were not taken into account.

Fig. 3 (January) and Fig. 4 (December) show two temporal slices from the video (no. 1), included to emphasise how the malaria risk continuously changes over the calendar year with the climate. The north-easterly winds from the Arabian Peninsula cause rainfall at the end of the year, which creates suitable climatic conditions for breeding and growth of mosquitoes (and development of the parasite), resulting in fuzzy values up to 0.27 for January and reaching 0.39 for December. The video clearly indicates that the climatic conditions for the period February-April are generally unsuitable for malaria transmission anywhere in Sudan as the fuzzy values for these months do not reach higher values than 0.07, 0.01 and 0.09, respectively. However, they peak up in May in the southern parts of

the South Gordofan and Blue Nile states (up to 0.46), while they simultaneously being very low (0-0.1) in other parts of the country, particularly in the northern, north-east and central regions of North Gordofan. Data for the period June-October show the highest rates of climatic conditions encountered (up to 0.9) in large parts of the country, which is thus suitable for the growth and proliferation of both mosquito and para-

site. Generally, the suitability increases at this time due to moist south-westerly winds covering most of Sudan resulting in increased rainfall and humidity.

August is the month with the largest areas of high fuzzy values, while November's climate conditions are unsuitable due to the dry north-easterly winds that begin to strengthen and push South at that time. By the end of December they already cover the entire country



Fig. 3. Malaria prediction map for January.



Fig. 4. Malaria prediction map for December from 2004-2010 (darker shades indicate areas more suitable for malaria transmission).

with fuzzy values between 0 and 0.13, except in the southern part, i.e. the Kassala state and part of the Red Sea state where conditions are just barely suitable.

The African MARA seasonal malaria map provides an opportunity to compare the present study with other studies conducted on the African continent. The model used explains that malaria transmission varies between different modes such as strongly seasonal to epidemic (1-3 months), seasonal and endemic (4-6 months), perennial and endemic (7-12 months) and periods without transmission (Adjuik et al., 1998). Although this map is based on the available long-term climate data rather than actual malaria data, it agrees with the present study concerning the association between the suitability of malaria in Sudan and the number of months with malaria transmission, i.e. the period May to October.

Video no. 2 provides the changing rates of malaria cases by state from January to December after rescaling to the 0-1 level. In general, the highest rates are seen in the states of Blue Nile, Sinnar and Al Gazirah with the trend declining in the north-western direction. The period from January to April showed lower malaria rates (0-0.71) in comparison to the other months, while the months of May, June and December show slightly higher rates (0-0.78). The highest values were recorded for the period from July to November corresponding to the autumn season, which is characterized by rainfall and high humidity.

Table 1. Non-parametric testing of malaria rates and malaria predictions.

Month	Mann-Whitney U	Wilcoxon W	P-value
January	13	338	0.00
February	112	437	0.03
March	112	437	0.03
April	116	441	0.04
May	147	472	0.26
June	168	493	0.59
July	165	285	0.53
August	153	273	0.34
September	173	293	0.69
October	145	470	0.24
November	116	441	0.04
December	112	437	0.03

Comparison between predicted and reported cases

Obviously, the transmission suitability (the prediction maps) and the actual number of malaria cases per unit of area (the case rate maps) should correlate. However, this correlation is not a direct one as the *Anopheles* mosquito needs nearly a month of appropriate weather conditions to complete its life cycle. This introduces a time lag resulting in the case map of malaria for a given month being associated with the prediction map for the previous month.

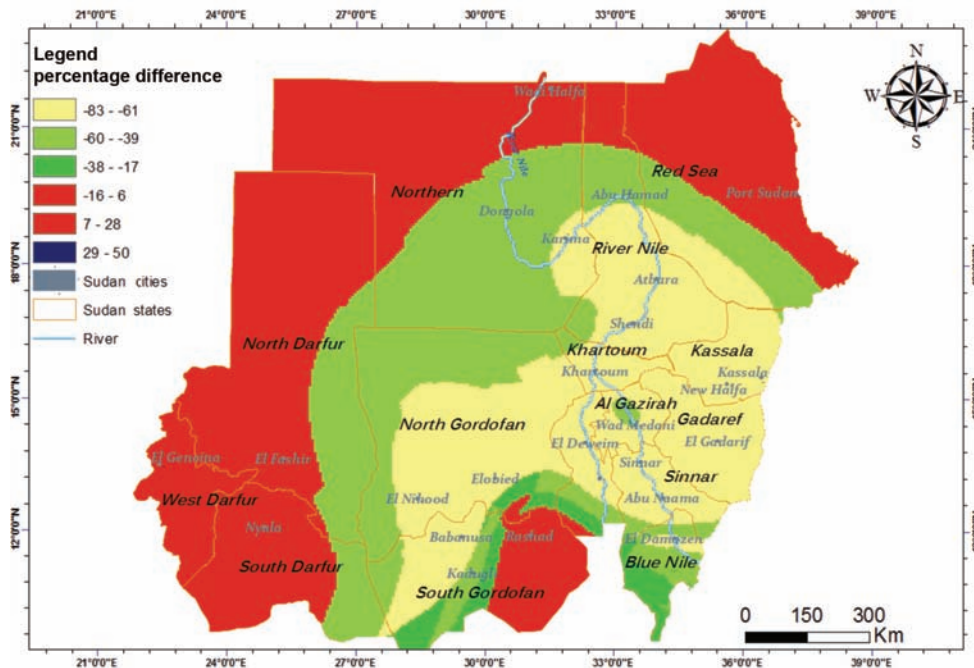


Fig. 5. Percentage difference between the prediction map of May and the case rate of June.

Berry (2007) used the *t*-test to compare two maps utilizing non-spatial, statistical tests for mapped data. He concluded that the *t*-test results are unable to verify the conditions of the test in terms of independent and normal distribution of the data. However, these conditions rarely hold for mapped data and the statistical tests ignore the spatial context of the data. In order to avoid the normality condition that is difficult

to achieve in spatial studies, we adopted a non-parametric approach. Table 1 shows the result of the Mann-Whitney *U* (Mann and Whitney, 1947) and the Wilcoxon *W* (Siegel, 1956), both non-parametric alternatives to the two-sample *t*-test used when comparing two related samples to assess groups of independent observations. The first group was the malaria rates and the second group the fuzzy sets. The null hypothesis

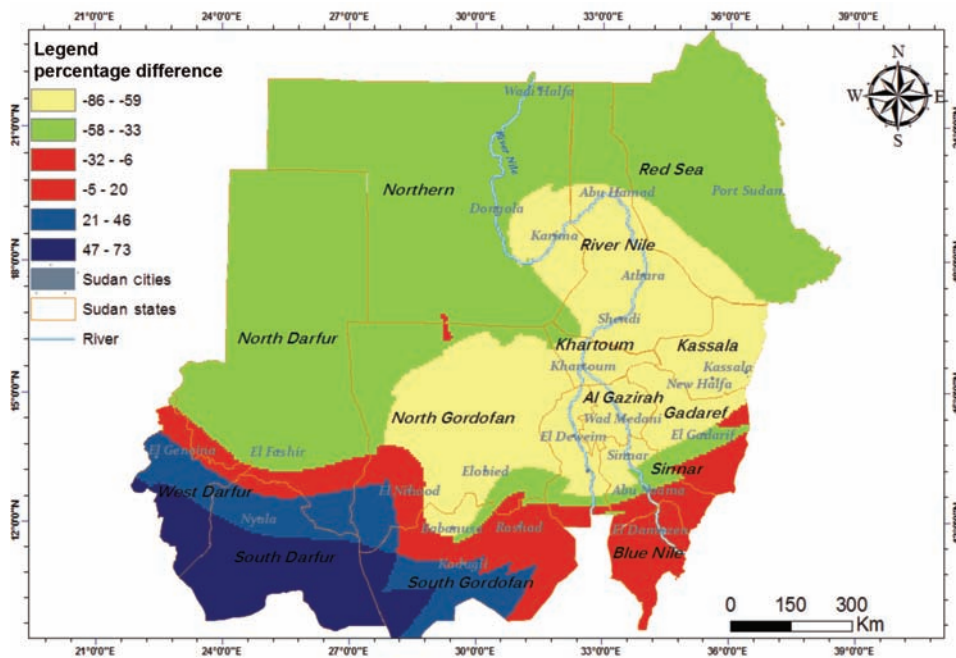


Fig. 6. Percentage difference between the prediction map of June and the case rate of July.

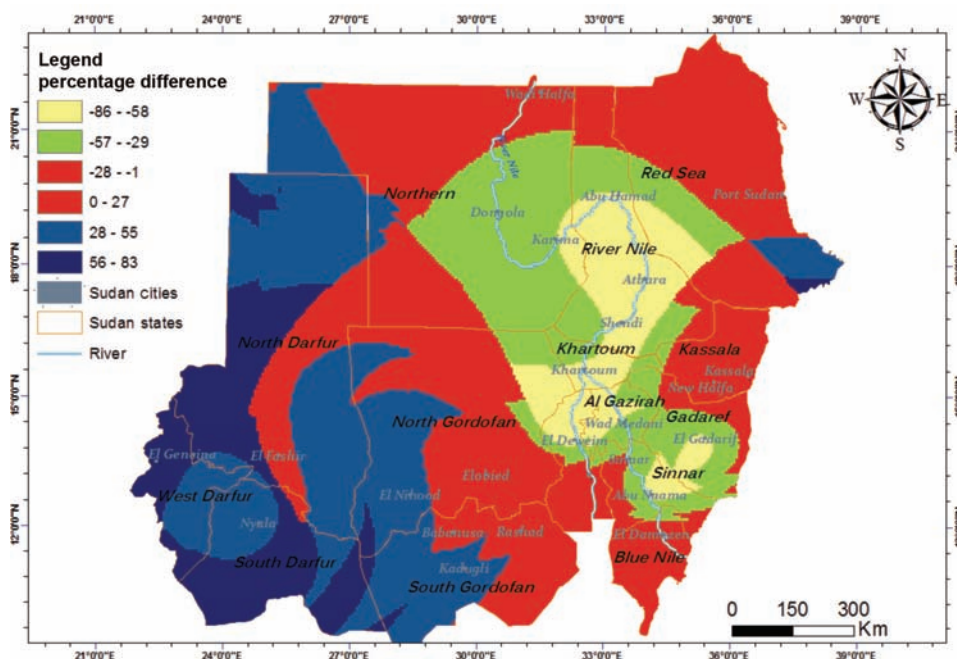


Fig. 7. Percentage difference between the prediction map of July and the case rate of August.

was that there should be no difference between the malaria rates and the malaria prediction values for any month. As can be seen in the table, the p-values for the months January, February, March, April, November and December were less than 0.05, which means that there actually were significant differences between the prediction values and those of the malaria rates. Hence, these months were omitted from the model, which then

compares May, June, July, August, September and October, for which the p-value equaled 0.26, 0.59, 0.53, 0.34, 0.69 and 0.24, respectively.

The percentage difference was used to compare prediction maps with the case rate maps. Fig. 5 to Fig. 10 show the percentage difference between two surfaces at each grid location. The dark-red areas along the maps correspond to areas with +/- 33% difference.

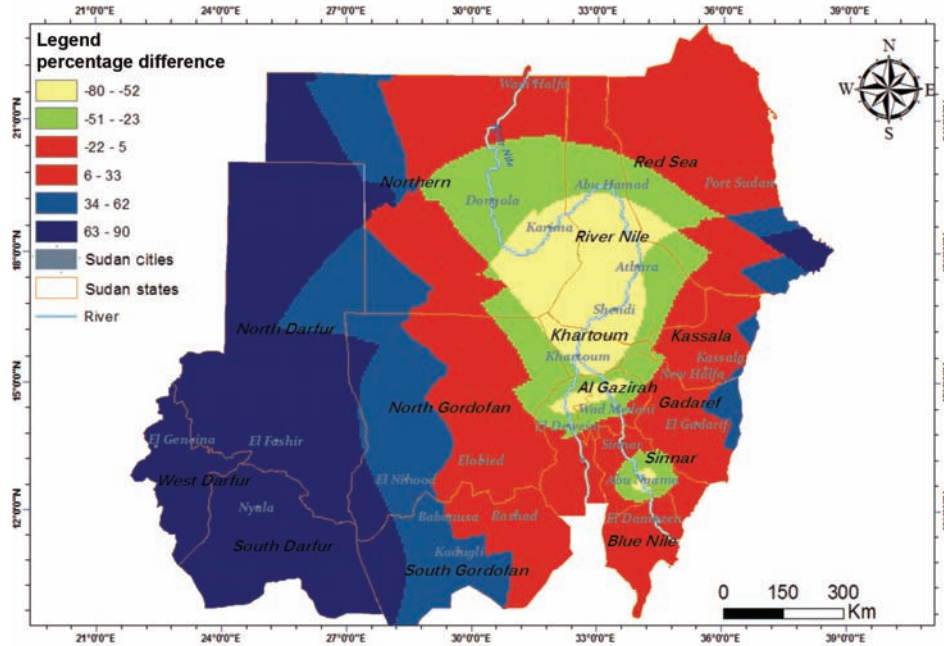


Fig. 8. Percentage difference between the prediction map of August and the case rate of September.

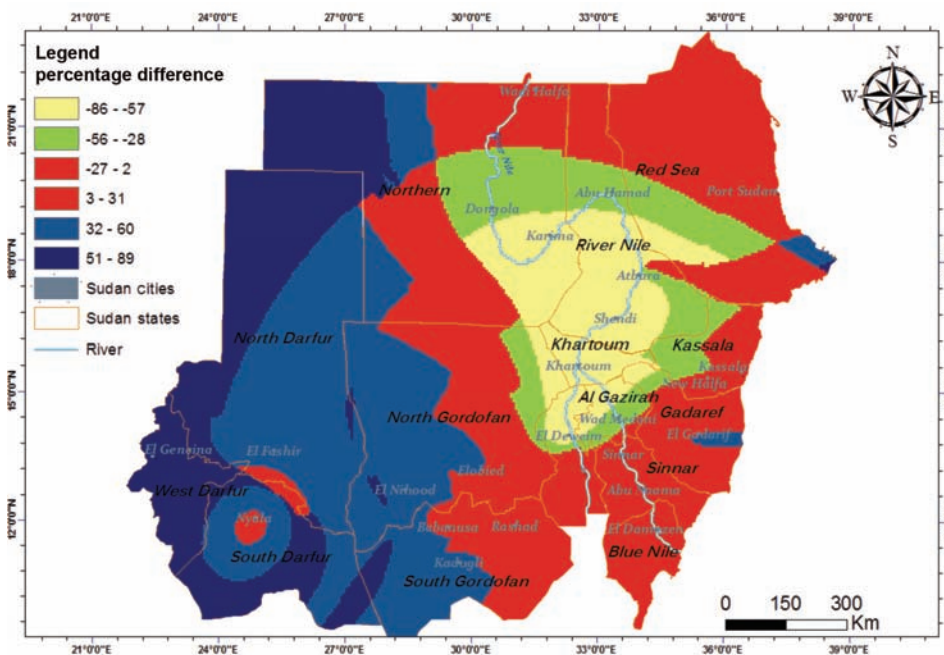


Fig. 9. Percentage difference between the prediction map of September and the case rate of October.

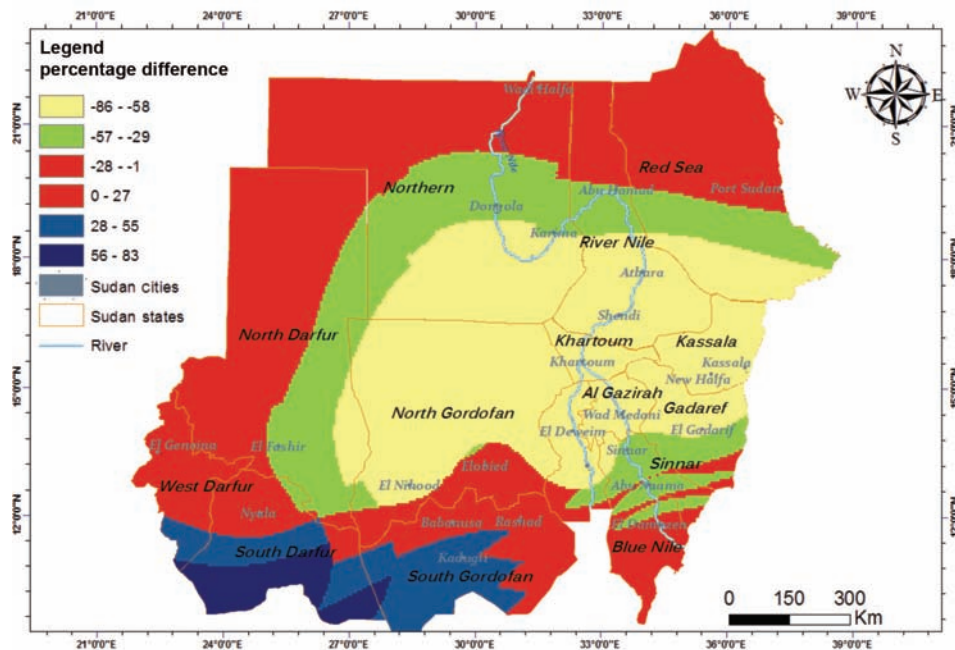


Fig. 10. Percentage difference the prediction map of October and the case rate of November.

According to Berry (2007) the rule for comparing map surfaces, if two-third of the map area is within one-third (33% difference), the surfaces are fairly similar; if less than one-third of the area is within one-third difference, the surfaces are fairly different. Fig. 5 shows that 51% of total area surface is within 33% difference when the prediction malaria map for May is compared to the case rate map for June. This means that less than two-third of the total surface area is within one-third difference indicating that the two maps are 50% similar. Fig. 6 shows the comparison between the prediction of malaria map in June and the case rate map for July. Here, 18% of total area surface is within the 33% difference indicating that the two maps are clearly different. In Fig. 7, around 47% of surface similarity can be seen between the prediction map for July and the case rate map for August, while 40% is within 33% difference of prediction malaria map in August and case rate map in September (Fig. 8). Fig. 9 shows that there is 43% similarity between prediction and case rate for September and October, respectively. Finally, the October prediction map and the November case rate map (Fig. 10) show that 55% of total area meets the similarity within 33% difference.

Acknowledgements

We are grateful to the Malaysian Technical Cooperation Program (MTCP) for financial support; many thanks to the

Economics and Social Research Bureau (ESRB), Ministry of Science and Technology, Sudan.

References

- Adjuik M, Bagayoko M, Binka F, Coetzee M, Cox J, Craig M, Deichman U, de Savigny DF, Fraser C, Gouws E, Kleinschmidt I, Lemardeley P, Lengeler C, Le Sueur D, Omumbo J, Snow B, Sharp B, Tanser F, Teusche T, Toure Y, 1998. Towards an atlas of malaria risk in Africa. First technical report of the mapping malaria risk in Africa/Atlas du risque de la malaria en Afrique (MARA/ARMA). Available at: http://www.mara.org.za/tr/ENG_MARA_Tech_Rep.pdf (accessed on)
- Berry JK, 2007. Map analysis: understanding spatial patterns and relationships. BASIS Press, University of Denver.
- Cox J, Craig M, Le Sueur D, Sharp B, 1999. Mapping malaria risk in the highlands of Africa. MARA/HIMAL Tech Rep 114.
- Craig M, 2009. The temporal and spatial distribution of malaria in Africa, with emphasis on Southern Africa. University of Basel, Switzerland.
- Craig M, Snow RW, le Sueur D, 1999. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today* 15, 105-111.
- Gillies MT, De Meillon B, 1968. The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). *South African Institute for Medical Research* 54, 1-343.
- Hardy JT, 2003. Climate change: causes, effects and solutions. Wiley, New York.
- Hutchinson MF, 1998. Interpolation of rainfall data with thin

- plate smoothing splines-part I: two dimensional smoothing of data with short range correlation. *J Geogr Inform Decis Anal* 2, 139-151.
- Hutchinson MF, Bischof RJ, 1983. A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter valley, New South Wales. *Aust Meteorol Mag* 31, 179-84.
- Hutchinson MF, McKenney DW, Lawrence K, Pedlar JH, Hopkinson RF, Milewska E, Papadopol P, 2009. Development and testing of Canada, wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961-2003. *J Appl Meteor Climatol* 48, 725-741.
- Hutchinson MF, Nix HA, McMahon JP, Ord KD, 1996. The development of a topographic and climate database for Africa. *Proceedings of the Third International Conference Integrating GIS and Environmental Modeling*. Santa Fe, New Mexico, USA: National Center for Geographic Information and Analysis.
- Jepson WF, Moutia A, Courtois C, 1947. The malaria problem in Mauritius: the bionomics of Mauritian *Anopheles*. *Bull Entomol Res* 38, 177-208.
- Mann HB, Whitney DR, 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* 18, 50-60.
- Martens WJ, Niessen LW, Rotmans J, Jetten TH, McMichael AJ, 1995. Potential impact of global climate change on malaria risk. *Environ Health Persp* 103, 458-464.
- McMichael AJ, 2003. *Climate change and human health: risks and responses*. Geneva: World Health Organization.
- Metz HC, 1992. *Sudan: a country study*. Washington DC, Federal Research Division.
- New M, Lister D, Hulme M, Makin I, 2002. A high-resolution data set of surface climate over global land areas. *Clim Res* 21, 1-25.
- Omumbo JA, Hay SI, Guerra CA, Snow RW, 2004. The relationship between the *Plasmodium falciparum* parasite ratio in childhood and climate estimates of malaria transmission in Kenya. *Malar J* 3, 17.
- Onori E, Grab B, 1980. Indicators for the forecasting of malaria epidemics. *Bull World Health Organ* 58, 91-98.
- Pampana E, 1969. *A textbook of malaria eradication*. Oxford University Press, London.
- Siegel S, 1956. *Non-parametric statistics for the behavioral sciences*. New York, McGraw-Hill, pp. 75-83.
- Wahba G, 1983. Bayesian "confidence intervals" for the cross-validated smoothing spline. *J Roy Stat Soc B Met* 45, 133-150.
- Zadeh LA, 1965. Fuzzy sets. *Inform Control* 8, 338-353.