Geospatial Health 2, 2007, pp. 213-222

The spatial distribution of *Anopheles gambiae sensu stricto* and *An. arabiensis* (Diptera: Culicidae) in Mali

N. Sogoba[1,2], P. Vounatsou[2], M.M. Bagayoko[3], S. Doumbia[1], G. Dolo[1], L. Gosoniu[2], S.F. Traore[1], Y.T. Toure[4], T. Smith[2]

[1]Malaria Research and Training Center, Faculté de Médecine, Pharmacie et Odontostomatologie, Université de Bamako, Mali; [2]Department of Public Health and Epidemiology, Swiss Tropical Institute, P.O. Box, CH-4002 Basel, Switzerland; [3]Vector Biology and Control Unit, Division of Prevention and Control of Communicable Diseases, WHO-AFRO, Gabon, BP 820, Libreville, Gabon; [4]Special Programme for Research and Training in Tropical Diseases (TDR), World Health Organization, CH-1211 Geneva, Switzerland

Abstract. Variations in the biology and ecology and the high level of genetic polymorphism of malaria vectors in Africa highlight the value of mapping their spatial distribution to enhance successful implementation of integrated vector management. The objective of this study was to collate data on the relative frequencies of *Anopheles gambiae s.s.* and *An. arabiensis* mosquitoes in Mali, to assess their association with climate and environmental covariates, and to produce maps of their spatial distribution. Bayesian geostatistical logistic regression models were fitted to identify environmental determinants of the relative frequencies of *An. gambiae s.s.* and *An. arabiensis* species and to produce smooth maps of their geographical distribution. The frequency of *An. arabiensis* was positively associated with the normalized difference vegetation index, the soil water storage index, the maximum temperature and the distance to water bodies. It was negatively associated with the minimum temperature and rainfall. The predicted map suggests that, in West Africa, *An. arabiensis* is concentrated in the drier savannah areas, while *An. gambiae s.s.* prefers the southern savannah and land along the rivers, particularly the inner delta of Niger. Because the insecticide knockdown resistance (*kdr*) gene is reported only in *An. gambiae s.s.* in Mali, the maps provide valuable information for vector control. They may also be useful for planning future implementation of malaria control by genetically manipulated mosquitoes.

Keywords: *Anopheles arabiensis, Anopheles gambiae s.s.*, Bayesian inference, geostatistics, kriging, malaria, Mali, Markov chain Monte Carlo.

Introduction

There are approximately 400 species of mosquitoes of the genus *Anopheles* (Culicidae) of which 30-40 transmit human malaria. In Africa, malaria transmission is mainly associated with *Anopheles gambiae sensu lato* (*An. gambiae s.l.*) and *An. funestus*. *An. gambiae s.l.* constitutes a complex of seven species with different abilities to transmit the parasite (White, 1974; Coluzzi et al., 1979; Coluzzi, 1984, 1994). In West Africa, the *An. gambiae* complex dominates, comprising mainly *An. gambiae s.s.* and *An. arabiensis* of which the former is itself undergoing a complicated process of incipient speciation. So far, five chromosomal (Bamako, Mopti, Savannah, Forest and Bissau) and two molecular (M and S) forms of *An. gambiae s.s.* have been identified (Coluzzi et al., 1985; Toure et al., 1998).

The species of *An. gambiae s.l.* and the genetic populations of *An. gambiae s.s.* vary in relative frequency, both seasonally and geographically. These

Corresponding author:
Penelope Vounatsou
Department of Public Health and Epidemiology
Swiss Tropical Institute
P.O. Box, CH-4002 Basel
Switzerland
Tel. +41 284 8109; Fax +41 284 8105
E-mail: Penelope.vounatsou@unibas.ch

remarkable differences in the biology and ecology and the high level of genetic polymorphism of the *An. gambiae s.l.* species highlight the value of mapping their spatial distribution to enhance effective implementation of integrated vector management (IVM) (Toure et al., 2004).

Maps of the spatial distribution of *An. gambiae s.l.* species have been produced by displaying the relative frequency of species at sampled locations (Toure et al., 1998; Coetzee et al., 2000; Onyabe and Conn, 2001), by climatic suitability conditions of the species (Lindsay et al., 1998) and by ecological niche-modeling (Levine et al., 2004). The latter links vector data with climatic factors using artificial-intelligence algorithms. However, only sparse data are available with which to build spatial distribution maps (Lindsay et al., 1998) and most of the predicted distribution maps currently available have been developed at the continental or sub-regional scale.

In Mali, the Malaria Research and Training Center (MRTC), University of Bamako, have gathered a countrywide dataset on *An. gambiae s.l.* species (*An. arabiensis* and *An. gambiae s.s.*) and sub-species (Bamako, Mopti and Savanna). We have now compiled, both published (Toure et al., 1998) and unpublished data from this database and used Bayesian geostatistical modeling to assess the spatial distribution of the two major vector species (*An. gambiae s.s.* and *An. arabiensis*) of *An. gambiae s.l.* in Mali. To our knowledge this is the first effort to produce maps of malaria vector species distribution adjusted for climatic factors using ground-truth data, and rigorous spatial statistical modeling at the country level.

Materials and methods

Description of the study area

The study area covers most of the territory of Mali in West Africa, i.e. a region between the latitudes 10° and 25° north and the longitudes 12° west and 4° east. Mali has an area of 1,240,000 Km^2 and a population estimated at 13,000,000 inhabitants in 2003 by the United Nations. The country is relatively flat, altitudinal variations are minimal, ranging from 200 to 350 m above sea level. The year is divided into two main seasons varying in length according to the latitude: a dry season (October–May) and a rainy season (June–September) characterized by lower temperatures and increased humidity.

Mali is drained by two major river systems (Senegal and Niger) and characterized by the following six eco-geographic strata:

(i) the southern Sudan savannah with an annual rain of 1300-1500 mm from May to October and a mean annual thermal amplitude of 5-6°C;
(ii) the northern Sudan savannah with about 700-1300 mm annual rainfall distributed over less than 6 months;
(iii) the Sahelian zones with 200-700 mm of annual rain distributed over three months and mean annual thermal amplitude of about 12°C;
(iv) the sub-Saharan zone with less than 200 mm of annual rain and 16°C of annual average thermal amplitude;
(v) the inner delta of the Niger River, a kind of "internal sea" between the northern Sudan savanna and the Sahelian zones, about 300 km long and 100 km wide, which influences the climate of the area, especially by reducing the average annual thermal amplitude; and
(vi) the Sahara desert where drought limits mosquito breeding.

Except in the most northerly part in the Sahara desert, the country is endemic for malaria (hyperendemic to hypoendemic when moving from South to North). The main malaria vectors are *An. gambiae s.l.* and *An. funestus*. *An. gambiae s.l.* is composed of *An. arabiensis* and three chromosomal forms of *An. gambiae s.s.* named Bamako, Mopti and Savannah (Toure et al., 1983).

Vector data

Both published (Toure et al., 1998) and unpublished data of the different research activities of the

MRTC, University of Bamako, Mali, were collated in a unique database. The data were obtained from cross-sectional and longitudinal surveys carried out between 1981 and 2004. Most surveys were conducted during the wet season (June-November). Survey sites were mainly small human settlements located in rural areas representing various eco-climatic zones of Mali. The database includes data collected from 94 locations and contains: (i) the total number of specimens; (ii) the count of *An. gambiae s.s.* and *An. arabiensis*; and (iii) the time of the survey (month and year). The specimens were differentiated by the chromosomal identification techniques (Coluzzi, 1968; Hunt, 1973) and/or by polymerase chain reaction (PCR) (Scott et al., 1993). The use of similar standardised techniques for sampling and processing mosquitoes across surveys ensured data consistency.

Climatic and environmental data

Factors used in this study were temperature, rainfall, the normalized difference vegetation index (NDVI), distance to water bodies, soil water storage (SWS) index, land use, agro-ecological zones and suitability to malaria transmission, a binary variable defined from environmental factors (Gemperli et al., 2006). A list of the data sources and spatial resolution is given in Table 1.

For each location, temperature and rainfall data were available as monthly long-term averages. NDVI data were also summarized by monthly long-term averages of the original decadal values during the period of 1985 to 1995. The agro-ecological zones (AEZ) were distinguished on the basis of the length of the growing period and were defined as follow: (i) the Equatorial Forest zone (>270 days); (ii) the Guinea Savannah zone (165-270 days); (iii) the Sudan Savannah zone (90-165 days); and (iv) the Sahelian zone (<90 days). In Mali only the last three AEZs can be found.

Data analysis

A buffer zone of 2 km around each data point was created using IDRISI 3.2 (Clark Labs, Clark University, MA, USA). The mean value of all pixels (with resolutions between 1 to 8 km^2 depending on the environmental factor) in this buffer area was calculated and used as the value of the given climatic and environmental factor. To take into account the possible lag time, between the rainfall and NDVI with the mosquito abundance, four summary measures (sum for rainfall and average for NDVI) were calculated for each one of the two climatic conditions:

(i) the mean climatic value during the month of collection (mean_1);
(ii) the mean climatic value during the previous month (mean_2);
(iii) the mean climatic value during the month of collection and the previous month (mean_3); and
(iv) the climatic value during the collection month and the two previous months (mean_4).

Vector data obtained from surveys extended over

Table 1. Climatic data sources and spatial resolution used in the study.

Factor	Spatial resolution	Source
Temperature	5 km^2	Hutchinson et al., 1996
Rainfall	5 km^2	Hutchinson et al., 1996
NDVI	8 km^2	NASA-AVHRR Land data sets (Agbu and James, 1994)
Land use	1 km^2	USGS-NASA
Water bodies	1 km^2	African Data Sampler World Resources Institute, 1995
SWS index	5 km^2	Droogers et al., 2001
Agro-ecological zone	Vector coverage	FAO, 1978

a period longer than a month were available, but cumulative for the whole period instead of monthly. In this case the midpoint month was used to relate the climatic factors.

Bivariate logistic regression models were fitted in STATA 9.0 (Stata Corporation, USA) to assess the relation between the proportion of *An. gambiae s.l.* vectors identified as *An. arabiensis* and the climatic and environmental factors. The Akaike's information criterion (AIC) was used to select the best summary measure and lag time for the rainfall and NDVI. The statistical significance of the environmental factors was assessed using the likelihood ratio test (LRT). All factors significant at the 15% significance level were entered into a Bayesian geostatistical multiple logistic regression model. The model took into account spatial heterogeneity by including the location-specific random effects ϕ_i at the sampling location level. In particular, we assumed that the *An. arabiensis* frequency Y_i at the sampling location *i* follows a binomial distribution, that is $Y_i \sim Bn\ (p_i,\ N_i)$, where N_i corresponds to the total number of *An. arabiensies* and *An. gambiae s.s.* mosquitoes collected, and p_i represents the *An. arabiensis* proportion at the location *i*. We further assumed that ϕ_i models a latent spatial process, that is $\phi = (\phi_i\ ,\ldots\phi_N)^T \sim MVN\ (0,\sum)$, with the covariance matrix $\sum$ a function of distance between locations, irrespective of the locations themselves (stationarity) and of the direction (isotropy). We adopted an exponential correlation function, that is $\sum_{ij} = \sigma^2 \exp(-\rho d_{ij})$ where σ^2 is the spatial variance, ρ the parameter that models the rate of correlation decay, and d_{ij} the distance between the locations *i* and *j*. Based on the above specification, the minimum distance for which the spatial correlation becomes less than 5% is calculated by $3/\rho$ (Ecker and Gelfand, 1997). The model parameters were estimated using Markov chain Monte Carlo (MCMC) simulation methods.

Bayesian kriging was used to predict the species frequency at 85,000 locations that were not sampled (Diggle and Tawn, 1998). The Bayesian model fit was carried out in WinBUGS 1.4. (Spiegelhalter et al., 2004), whereas the model prediction was implemented in Fortran 95 (Compaq Visual Fortran, Professional 6.6.0) using standard numerical libraries (NAG, The Numerical Algorithms Group Ltd).

Results

The results of the bivariate logistic regression analyses are shown in Table 2 which indicates that, among the four NDVI and rainfall measures considered in the study, the ones which fitted the *An. arabiensis* proportion best (giving smaller AIC) were the NDVI mean value during month of collection and the sum of rainfall mean value during month of collection and the two previous months, respectively. The bivariate analyses also revealed that the agro-ecological zone, distance to water bodies, land use, transmission suitability, SWS index, minimum and maximum temperature were significantly associated with the relative frequency of *An. arabiensis*, which increases from the Guinea to the Sahelian AEZ. The crop/grass land mosaic and water body categories of land use, the minimum temperature and the suitability to the transmission were negatively associated with the *An. arabiensis* frequency at a significant level.

All the factors above were entered into a Bayesian geostatistical model. The results of the spatial multiple regression model showed that the sum of the mean rainfall during collection month and the two previous months and the minimum temperatures were the only factors negatively associated with the relative frequency of *An. arabiensis*. None of the land use categories were significantly related to the proportional presence of this mosquito strain. Comparing the different categories of the variables between the bivariate and the multiple regression models, the following changes were observed: the Sahel category of the AEZ and the crop/grass land mosaic and water body categories of land use changed from significant in the bivariate model to not significant in the multiple regression model; the 4-10 km distance category of the distance to water bodies, the NDVI mean value

Table 2. Bivariate and multiple spatial logistic regression models of *An. arabiensis* relative frequency with climate and environmental variables.

	Bivariate analysis				Spatial model	
Variables	Lag*	Coefficient	95% CI	p-value (AIC)	Median	95% CI
Agro-ecological zones						
Guinea savannah	-	0.00			0.00	
Sudan savannah	-	2.16	2.07, 2.25	<0.001	2.01	0.25, 3.49
Sahel	-	2.72	2.55, 2.89		2.49	-1.21, 5.54
Distance to water bodies (km)						
< 4	-	0.00			0.00	
4 - 10	-	-0.29	-0.38, -0.20	<0.001	1.58	0.65, 2.58
>10 - 20	-	0.82	0.77, 0.87		1.51	0.52, 2.50
> 20	-	1.77	1.65, 1.89		2.02	0.77, 3.36
Land use categories						
Savannah	-	0.00			0.00	
Crop/grass land/mosaic	-	-2.07	-2.38, -1.77		-0.65	-2.80, 1.53
Grass land	-	-0.02	-0.09, 0.06		0.51	-1.58, 2.91
Shrub land	-	-0.39	-1.02, 0.24	<0.001	-2.50	-7.05, 1.69
Water bodies	-	-0.95	-1.72, -0.18		-0.19	-2.91, 2.46
Barren/sparsely vegetated	-	0.23	-0.14, 0.60		-1.24	-5.58, 2.32
Suitability with regard to transmission						
Not suitable	-	0.00			0.00	
Suitable	-	-0.15	-0.21, -0.08	< 0.001	0.10	-0.02, 0.22
Rainfall						
Mean_1	0	-0.0001	-0.0003, 0.0001	0.573 (50410.8)	-	
Mean_2	1	0.0003	0.0000, 0.0005	0.0198 (50405.7)	-	
Mean_3	1	0.0000	-0.0001, 0.0002	0.439 (50410.5)	-	
Mean_4	2	0.0001	0.0000, 0.0002	0.004 (50403.0)	-0.01	-0.006, -0.004
NDVI						
Mean_1	0	-0.32	-0.49, -0.16	0.0001 (50396.6)	-	
Mean_2	1	-0.08	-0.25, 0.08	0.3285 (50410.6)	-	
Mean_3	1	-0.22	-0.39, -0.05	0.0129 (50404.9)	-	
Mean_4	2	-0.33	-0.52, -0.150	0.0005 (50398.9)	1058.0	8.67, 12.40
Temperature						
Mean minimum	-	-0.0030	-0.004, -0.002	<0.001	-0.02	-0.042, -0.004
Mean maximum	-	-0.0070	-0.008, 0.006	<0.001	0.21	0.18, 0.24
SWS index	-	0.22	0.13, 0.31	<0.001	1.71	1.43, 2.01
Spatial parameters						
ρ	-				4.00	2.63, 4.56
σ^2	-				0.04	0.02, 0.06

Legend. Mean_1 = climatic mean value during month of collection; Mean_2 = climatic mean value during the previous month; Mean_3 = climatic mean value during month of collection and the previous month; Mean_4 = climatic mean value during collection month and the 2 previous months; * Lag time (in month) between the environmental variables and the collection date (month) of vector data.

during the month of collection, and the two previous months (the one included in the multiple regression model) changed from negatively significant in the bivariate model to positively significant in the multiple regression model; the mean maximum temperature which was not significant in the bivariate model became positively significant in the multiple regression model.

Assuming that spatial correlation is a function of distance between locations, irrespective of the locations themselves (stationary) and of the direction (isotropic), the minimum distance at which that correlation was less than 5% was as much as 1333.4 km (95% CI = 913.4-1520.1).

Fig. 1 shows the observed relative frequencies of *An. arabiensis* and *An. gambiae s.s.* in the 94 locations. A lower frequency of *An. arabiensis* was observed in the southern and northern savannah while higher frequencies were observed in the Sahelian zone, with the exception of the inner delta of Niger. Maps of the predicted proportions of *An. arabiensis* are shown in Fig. 2 which depicts a south to north distribution pattern of *An. arabiensis* relative frequency with a moderate proportion of *An. ara-*

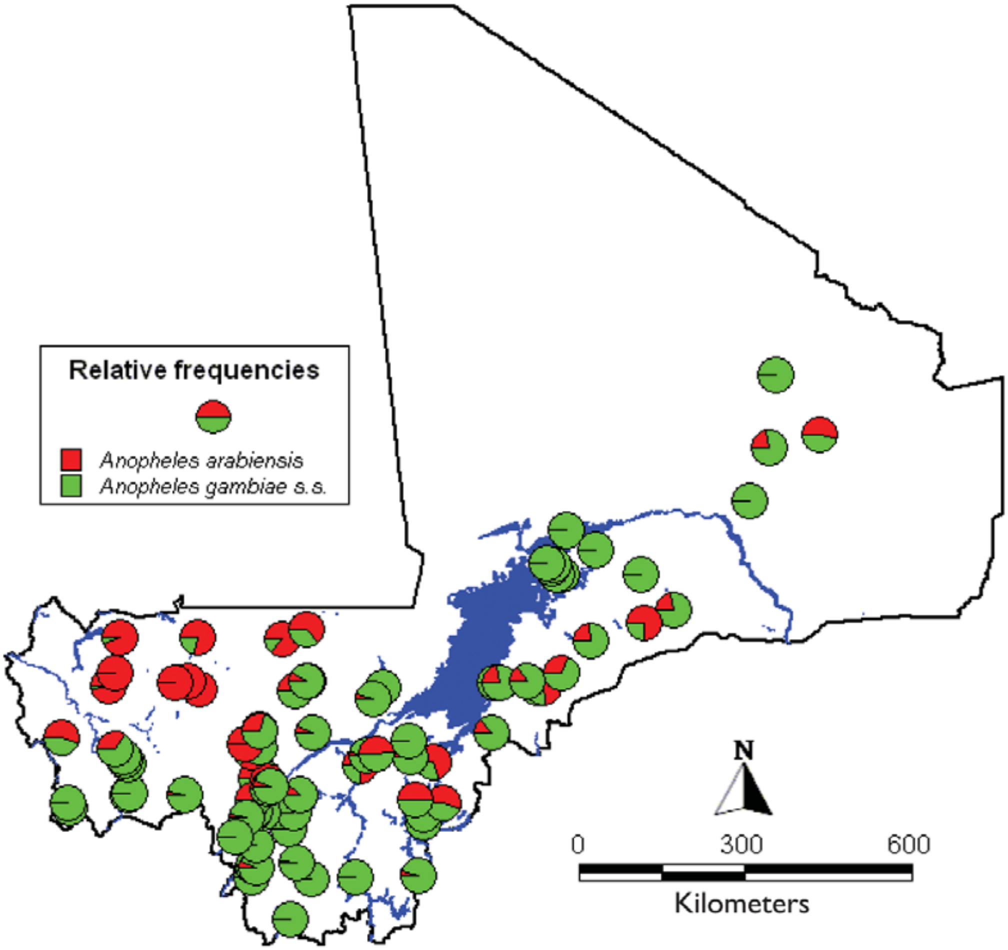

Fig. 1. Observed relative frequencies of *An. arabiensis* and *An. gambiae s.s.* in 94 sampling locations in Mali, West Africa. The green color represents the relative frequencies of *An. gambiae s.s.* and the red the relative frequencies of *An. arabiensis*.

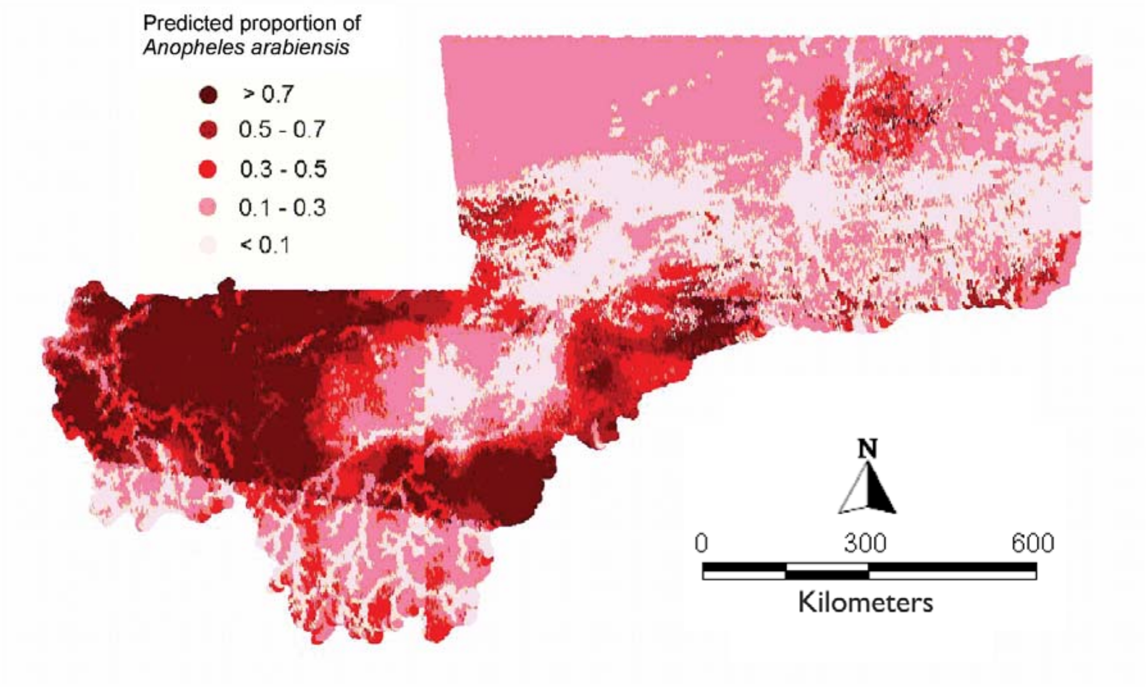

Fig. 2. Map of predicted relative frequencies of *An. arabiensis*.

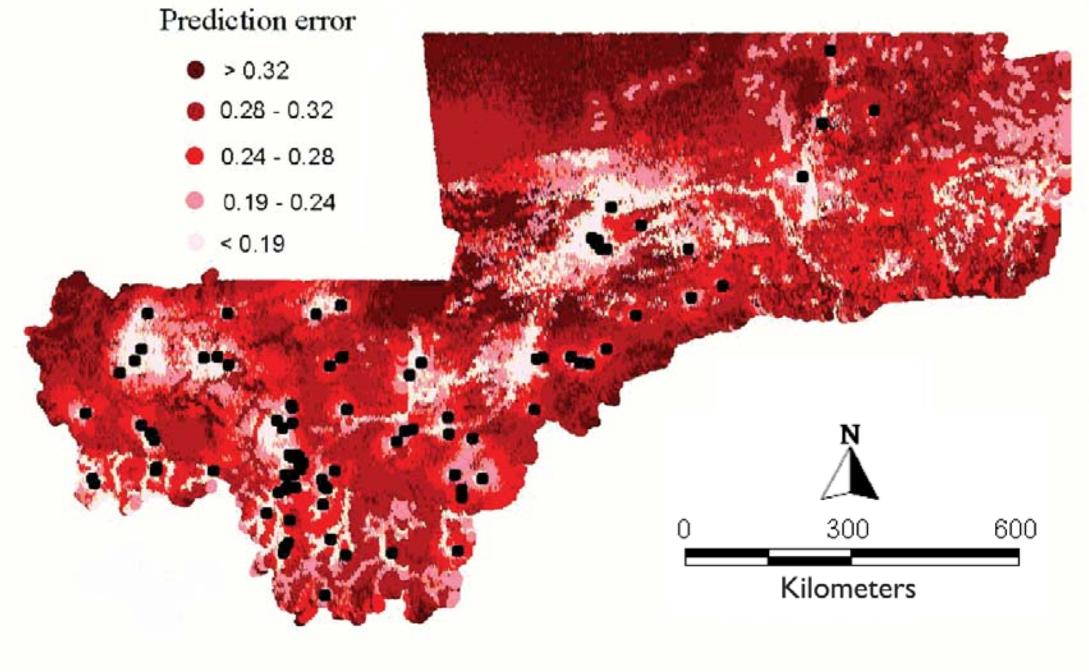

Fig. 3. Map of prediction error of the relative frequencies of *An. arabiensis*.

biensis in the southern savannah, a higher proportion in the northern savannah and Sahelian zones (apart from the inner delta of the Niger river where *An. arabiensis* was almost absent) and a lower one in the sub-Sahara zone. The *An. arabiensis* proportion is also lower along the rivers irrespective of the eco-climatic zone. Estimates of the prediction error are shown in Fig. 3. The prediction error is lowest along the rivers and increases with the distance from water bodies. In contrast, the prediction error is relatively high in the sub-Sahara zone where very few surveys were carried out.

Discussion

In this study, we compiled published and unpublished vector data in a unique database and using Bayesian geostatistical modeling, identified climatic and environmental factors associated with the relative frequency of the two major malaria mosquito vector species (*An. gambiae s.s.* and *An. arabiensis*) of *An. gambiae s.l.* in Mali, and assessed their spatial distribution. We used an approach considering different measures of rainfall and NDVI and performed bivariate logistic regressions to select the measures which fitted the data best using the AIC criterion. This was done to select the subset of variables to be fitted into the spatial model because Bayesian variable selection is not straightforward and requires specialized software which is not currently available. The approach adopted has been used also in other applications of spatial Bayesian modeling (Gemperli et al., 2006; Gosoniu et al., 2006). The results show that the cumulated rainfall value during the survey and during the two previous months, and the NDVI value during the survey month, fitted the data better than the other rainfall and NDVI measures assessed. This suggests that the *An. gambiae* complex species composition is more sensitive to the cumulated rainfall over previous months than to the value during the survey month. The observed lag time period between rainfall and vector abundance can enhance operational malaria early-warning systems (MEWS) based on rainfall estimates (Grover-Kopec et al., 2005; Thomson et al., 2006).

The two sibling species of the *An. gambiae* complex (*An. arabiensis* and *An. gambiae s.s.*) exist across the whole study area. The estimates of the spatial model for the proportion of *An. arabiensis* showed a positive association between the NDVI values, the SWS index, the maximum temperature, and the distance to the water bodies. Minimum temperature and rainfall were negatively related to the relative frequencies of *An. arabiensis*. The predicted map in Fig. 2 represents the median relative frequency of *An. arabiensis* over the transmission period (June to November). This is broadly in agreement with the ecological distribution of *An. arabiensis* in Mali (Toure et al., 1998). *An. arabiensis* is concentrated in the drier savannah areas and *An. gambiae s.s.* in the inner delta of Niger, the southern savannah and along the rivers. The occurrence of *An. arabiensis* in the drier savannah reflects the known preference of this species for drier conditions. The occurrence of *An. gambiae s.s.* in the arid regions (Sahel) has been shown to be associated with the ‘Mopti’ chromosomal form (Toure et al., 1994). Many studies across Africa have described the likely adaptation of *An. arabiensis* to drier conditions than *An. gambiae s.s.* (Coetzee et al., 2000; Onyabe and Conn, 2001; Kirby and Lindsay, 2004; Levine et al., 2004). The general association of this mosquito strain with river systems is illustrated by its positive association with the SWS index and NDVI. Laboratory and field experimentation also showed that *An. arabiensis* adults are better adapted to hotter conditions than *An. gambiae s.s.* (Robert, 1998; Kirby and Lindsay, 2004). The ability for *An. arabiensis* to withstand the dry season may explain the weak and negative association of *An. arabiensis* relative frequency with rainfall.

The same pattern of south to north distribution of *An. arabiensis* relative frequencies was observed with the transmission model (Gemperli et al., 2006). However, in contrast to the distribution of *An. arabiensis*, the transmission model showed higher entomological inoculation rate in the south and moderate to low in the middle and northern part of country. This suggests that *An. arabiensis* may contribute less to the transmission than *An. gambiae s.s.*

Fig. 2 depicts the spatial distribution of *An. arabiensis* and *An. gambiae s.s.* over the whole transmission period. Other studies (White, 1974; Coluzzi et al., 1979; Coluzzi, 1984, 1994) have shown that the temporal distribution is one of the key elements in malaria epidemiology and vector control which has valuable implication for vector stratification and adequate planning of both vector control and research activities. Our study did not take into account temporal aspects for two reasons: firstly temporally disaggregated environmental data were not available for all survey years, especially not for the surveys conducted in the early 1980s; secondly the vector data were generally reported pooled from several surveys. Nevertheless, our effort was to produce predicted maps of the spatial distribution of *An. arabiensis* and *An. gambiae s.s.* species adjusted for climatic factors using spatial statistical modeling supported by consistently observed vector data. The advantage of our study over preceding ones is that we used statistical analysis which quantifies the relationship between environment-vector data and identifies significant determinants instead of only using geographical information system. The Bayesian approach we used takes into account the spatial dependence present in the data in a flexible way and calculates inherently the standard errors of the parameter estimates as well as the prediction error without relying on approximations or asymptotic results. The map of the prediction error indicates the confidence we can have on the model prediction for the study area.

A practical implication of our findings is their relevance in monitoring of insecticide resistance encoded by the *kdr* gene. In Mali resistant alleles of *kdr* have been reported only in the chromosomal form Savannah of *An. gambiae s.s.* (Fanello et al., 2003). Based on these results, insecticide resistance monitoring and management must be primarily focused on the humid savannah, along the rivers and in the inner delta of Niger where a higher frequency of *An. gambiae s.s.* is encountered. Understanding the spatial distribution of *An. gambiae s.l.* species and sub-species may also be a prelude to a successful implementation of genetic control, such as the use of transgenic technologies to make mosquitoes refractory to the parasite. IVM strategies that target particular vector populations will need information at high spatial and temporal resolutions on the distribution of the sibling species of *An. gambiae* complex.

Acknowledgements

The authors would like to acknowledge Olivier Briët for helping in the environmental data extraction. We are thankful to Drs. Richard Sakaï and Robert Gwadz for their support. The analysis of the data was supported by the Swiss National Science Foundation project Nr. 3252B0-102136/1.

References

Agbu PA, James ME, 1994. NOAA/NASA pathfinder AVHRR land data set user's manual. Goddard distributed active archive center, NASA Goddard Space Flight Center.

Coetzee M, Craig M, le Sueur D, 2000. Distribution of African malaria mosquitoes belonging to the *Anopheles gambiae* complex. Parasitol Today 16, 74-77.

Coluzzi M, 1968. Cromosomi politenici delle cellule nutrici ovariche nel complesso *gambiae* del genere *Anopheles*. Parassitologia 10, 179-183.

Coluzzi M, 1984. Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control. Bull World Health Organ 62, 107-113.

Coluzzi M, 1994. Malaria and the Afrotropical ecosystems: impact of man-made environmental changes. Parassitologia 36, 223-227.

Coluzzi M, Petrarca V, Di Deco MA, 1985. Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. Boll Zool 52, 45-63.

Coluzzi M, Sabatini A, Petrarca V, Di Deco MA, 1979. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans R

Soc Trop Med Hyg 73, 483-497.

Diggle PJ, Tawn JA, 1998. Model-based geostatistics. Appl Stat 47, 299-350.

Droogers P, Seckler D, Makin I, 2001. Estimating the potential of rainfed agriculture. International Water Management Institute Working, Paper 20.

FAO, 1978. Report on the agro-ecological zones project, Vol. 1, Methodology and results for Africa. World Soil Resources Report 48, 32-41.

Ecker M, Gelfand AE, 1997. Bayesian variogram modelling for an isotropic spatial process. JABS 4, 347-369.

Fanello C, Petrarca V, della TA, Santolamazza F, Dolo G, Coulibaly M, Alloueche A, Curtis CF, Toure YT, Coluzzi M, 2003. The pyrethroid knock-down resistance gene in the *Anopheles gambiae* complex in Mali and further indication of incipient speciation within *An. gambiae* s.s. Insect Mol Biol 12, 241-245.

Gemperli A, Vounatsou P, Sogoba N, Smith T, 2006. Malaria mapping using transmission models: application to survey data from Mali. Am J Epidemiol 163, 289-297.

Gosoniu L, Vounatsou P, Sogoba N, Smith T, 2006. Bayesian modelling of geostatistical malaria risk data. Geospatial Health 1, 127-139.

Grover-Kopec E, Kawano M, Klaver RW, Blumenthal B, Ceccato P, Connor SJ, 2005. An online operational rainfall-monitoring resource for epidemic malaria early warning systems in Africa. Malar J 4, 6.

Hunt RH, 1973. A cytological technique for the study of *Anopheles gambiae* complex. Parassitologia 15, 137-139.

Hutchinson MF, Nix HA, McMahon JP, Ord KD, 1996. Africa - a topographic and climate database (CD-ROM). The Australian National University.

Kirby MJ, Lindsay SW, 2004. Responses of adult mosquitoes of two sibling species, *Anopheles arabiensis* and *A. gambiae s.s.* (Diptera: Culicidae), to high temperatures. Bull Entomol Res 94, 441-448.

Levine RS, Peterson AT, Benedict MQ, 2004. Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm. Am J Trop Med Hyg 70, 105-109.

Lindsay SW, Parson L, Thomas CJ, 1998. Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae sensu stricto* and *An. arabiensis*, using climate data. Proc Biol Sci 265, 847-854.

Onyabe DY, Conn JE, 2001. The distribution of two major malaria vectors, *Anopheles gambiae* and *Anopheles arabiensis*, in Nigeria. Mem Inst Oswaldo Cruz 96, 1081-1084.

Robert V, 1998. Age grading *Anopheles arabiensis*: their gorging and surviving responses using a membrane feeding system. Parasite 5, 87-90.

Scott JA, Brogdon WG, Collins FH, 1993. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. Am J Trop Med Hyg 49, 520-529.

Spiegelhalter D, Thomas A, Best NG, Lunn D, 2004. WinBUGS users manual. Version 1.4.1.

Thomson MC, Doblas-Reyes FJ, Mason SJ, Hagedorn R, Connor SJ, Phindela T, Morse AP, Palmer TN, 2006. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. Nature 439, 576-579.

Toure YT, Oduola AM, Morel CM, 2004. The *Anopheles gambiae* genome: next steps for malaria vector control. Trends Parasitol 20, 142-149.

Toure YT, Petrarca V, Coluzzi M, 1983. Nuove entità del complesso *Anopheles gambiae* in Mali. Parassitologia 25, 367-370.

Toure YT, Petrarca V, Traore SF, Coulibaly A, Maiga HM, Sankare O, Sow M, Di Deco MA, Coluzzi M, 1994. Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae s.str.* in Mali, west Africa. Genetica 94, 213-223.

Toure YT, Petrarca V, Traore SF, Coulibaly A, Maiga HM, Sankare O, Sow M, Di Deco MA, Coluzzi M, 1998. The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West Africa. Parassitologia 40, 477-511.

White GB, 1974. *Anopheles gambiae* complex and disease transmission in Africa. Trans R Soc Trop Med Hyg 68, 278-301.

World Resources Institute, 1995. African Data Sampler, (CD-ROM) Edition I.