

Institut Sénégalais de Recherches Agricoles



Fiche Technique

**Ecological niche modelling to estimate distribution of
Culicoides, potential vectors of bluetongue virus in Senegal**

**Mamadou Ciss¹, Biram Biteye¹, Assane Gueye Fall¹, Moussa Fall¹, Marie
Cicille Ba Gahn^{1,2}, Louise Leroux^{3,4}, Andrea Apolloni^{5,6}**

Document presented by:

Mamadou Ciss¹, Biram Biteye¹, Assane Gueye Fall¹, Moussa Fall¹, Marie Cicille Ba Gahn^{1,2}, Louise Leroux^{3,4}, Andrea Apolloni^{5,6}



Institut Sénégalais de Recherches Agricoles/Laboratoire National de l'Élevage et de Recherches Vétérinaires BP 2057 Dakar-Hann, Sénégal.



²Université Cheikh Anta Diop, Faculté des Sciences et Techniques, Département de Biologie Animale, Laboratoire d'Écologie Vectorielle et Parasitaire, Dakar, Sénégal.



cirad CIRAD, UPR AIDA, Dakar, Sénégal. ⁴AIDA, Univ Montpellier, CIRAD, Montpellier, France.



cirad CIRAD, UMR ASTRE, Dakar, Sénégal. ⁶ASTRE, Univ Montpellier, CIRAD, INRA, Montpellier, France.

Background:

Vector-borne diseases are among the leading causes of morbidity and mortality in humans and animals. In the Afrotropical region, some of these are transmitted by culicoides, such as Akabane virus, bluetongue, epizootic hemorrhagic fever and African horse sickness [1-3]. Bluetongue virus is one of the most important livestock pathogens, capable of infecting domestic and wild ruminants, causing serious clinical disease in some and reducing productivity in other.

Method:

1. Study area

As a part of the nationwide surveillance program in Senegal in 2012, 108 sites holding livestock were initially selected (as trap-site) as follows: 3 sites per department and 3 departments per region among 12 of the 14 Senegalese regions. The sites were chosen in order to better understand the complex relationship between hosts-pathogens-vectors and the part of environmental and climatic factors on the emergence and/or re-emergence of vector-borne diseases such as BT and AHS. Ziguinchor and Sédhiou regions were excluded because of safety reasons. In this study we considered only data from 96 sites (Fig.1) that were visited at the end of the rainy season 2012 (between September and October).

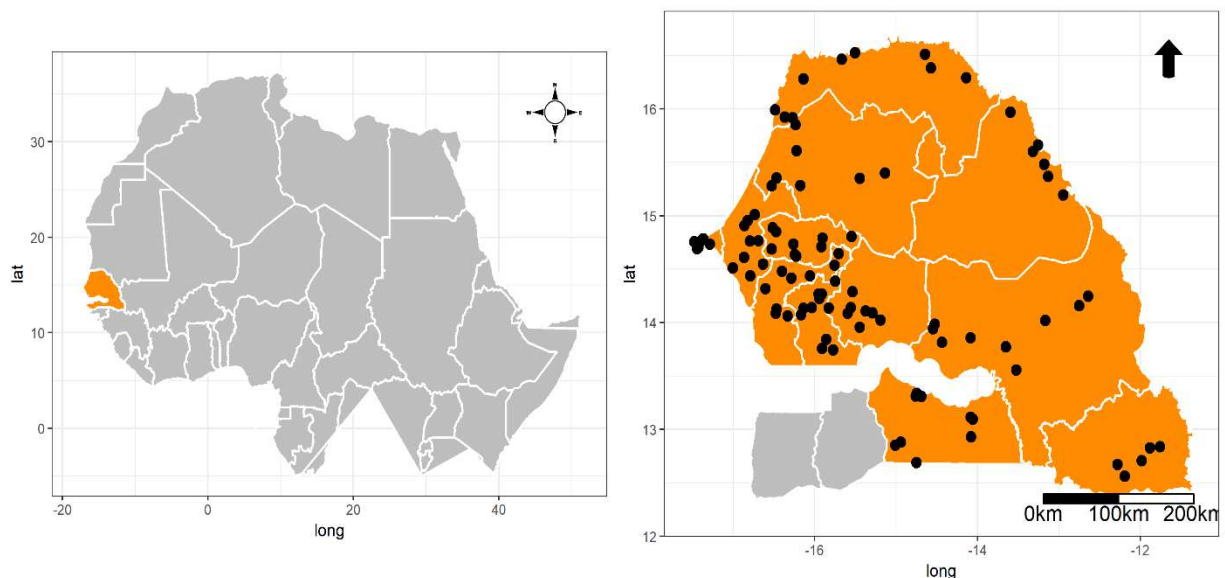


Figure 1: Map of Senegal, a western Africa country (A) with location of study sites in 12 Senegalese regions (B). In yellow, the study area and in grey, the no visited area.

2. Data collection

2.1. Entomological data

To collect information on the spatial distribution of culicoides species in Senegal, a nationwide culicoides trapping campaign was organized at the end of the 2012 rainy season (between September and October). Overall, with the help of the veterinary services, 108 sites were initially selected as follows: 3 sites per department and 3 departments per region in 12 (of the 14) regions of Senegal (Fig. 1). Culicoides specimens were collected during two consecutive nights in each site using Onderstepoort black-light suction traps (Onderstepoort Veterinary Institute South Africa) positioned close to livestock's. Morphological identification of Culicoides species was conducted by examination of the wing pigmentation pattern using a stereomicroscope as explained by [4]. For species that are difficult to identify, the specimens were dissected and slide-mounted in accordance with the Wirth and Marston technique for observation under a microscope. Several identification keys were used depending on the species found and their subgenus or group [5-8]. The geographical coordinates of each site were recorded with a Garmin© hand-held global positioning system receivers (accurate to within 10 meters) and projected in UTM Zone 28 N.

2.2 Climatic, environmental and livestock parameters

Several variables (26 in total) were used to implement the model. These were grouped in 4 categories (Table 1): 11 Bio-climatic variables related to temperature (Bio01 – Bio11); 8 Bio-climatic variables relative to precipitations (Bio12 – Bio19); Elevation data (1 variable) and Animal density (6 variables). Elevation data (digital elevation model) via the global elevation data (MOD13Q1), from the Moderate Resolution Imaging Spectroradiometer (MODIS) with a spatial resolution of 30 arc-seconds (~1km) have been downloaded from the Pale-Blu data archive website (<https://www.edenextdata.com/?q=content/modis-1km-digital-elevation-model-and-landwater-mask-v5>). The bioclimatic data, with a spatial resolution of 30 arc-seconds (~1km), have been downloaded from the World climate [9] website (<http://www.worldclim.org/current>) and averaged over a 50-year time period between 1950 and 2000 at a spatial resolution of 30 arc-seconds (~1km). Livestock data were used linking to their importance for vectors [10,11]. These livestock data (number of head/department) came from an unpublished study by the Direction des Services Vétérinaires (DSV), the Senegalese national institute competent body for animal health (DSV, 2013, unpublished work). Before stacked, livestock data (6 variables) were rasterized according to department with a spatial

resolution of 250m x 250m. All layers were resampled at a resolution 250m x 250m and were stacked in one spatial layer.

Table 1: Different variables and their abbreviations (code) used in the ENFA and Maxent niche models

Category of variables	Parameter	Abbreviation/Code
Temperature	Annual Mean Temperature	Bio01
	Mean Diurnal Range (Mean of monthly (max temp – min temp))	Bio02
	Isothermality (Bio02/Bio07)*100	Bio03
	Temperature seasonality (standard deviation *100)	Bio04
	Maximum Temperature of Warmest Month	Bio05
	Minimum Temperature of Coldest Month	Bio06
	Temperature Annual Range (Bio05-Bio06)	Bio07
	Mean Temperature of Wettest Quarter	Bio08
	Mean Temperature of Driest Quarter	Bio09
	Mean Temperature of Warmest Quarter	Bio10
	Mean Temperature of Coldest Quarter	Bio11
Precipitations	Annual Precipitation	Bio12
	Precipitation of Wettest Month	Bio13
	Precipitation of Driest Month	Bio14
	Precipitation seasonality (coefficient of variation)	Bio15
	Precipitation of Wettest Quarter	Bio16
	Precipitation of Driest Quarter	Bio17
	Precipitation of Warmest Quarter	Bio18
Precipitation of Coldest Quarter	Bio19	
Altitude	Digital elevation model	Dem
Animal density	Cumulated density of horse, cattle, donkey, goat and sheep	Livestock
	Cattle density	Cattle
	Goat density	Goat
	Sheep density	Sheep
	Horse density	Horse
	Donkey density	Donkey

3. Modelling habitat suitability

At first, we considered an exploratory analysis to identify the variables determining the suitability of the habitat (that is the presence of the species). To this aim we used Ecological Niche Factor Analysis (ENFA) [12,13], a multivariate approach, similar to PCA, that requires only presence data for the species in order to explore relationship between occurrence and environmental data. The first axis (marginality axis) of the ENFA is a measure capturing the dimension in the ecological space in which the average conditions where the species lives differ from the global conditions; a large marginality value implies that the conditions where the species is found are “far” from the global environmental conditions. In contrast, the

second axis (specialization) is a measure of the narrowness of the niche (ratio of the multidimensional variances of the available to occupied spaces).

After variables have been chosen, we used MaxEnt approach [14-16] to model the culicoides presence. MaxEnt is a machine learning technique based, as the term indicates, on the principle of maximum entropy for a probability distribution, given constraints on its momenta:

$$\text{Entropy} = - \sum_i p_i \log p_i$$

where p_i corresponds to the specie's presence probability in a cell i . The method uses presence locations and environmental covariates for all cells in the study area. Data, used to inform the model, define moment constraints on the distribution, while covariates define the mean, variance etc. where species occur. The result consists on the estimation of the presence probability in each cell. Entomological data (for each species) were randomly partitioned into two subsets: 75% of the records used for training and building up the model, and 25% for testing model accuracy. 10,000 random background points were used to evaluate commission. This was done to test the model performance given by Area Under the Receiver Operating Characteristic Curve (AUROC); a plot of sensitivity against specificity which measures the ability of the model to discriminate between sites where a species is present ($y = 1$) against where it is absent ($y = 0$) [17-19]. AUROC values range from 0 to 1; an AUROC value higher than 0.8 indicates robust performance of the model. Statistical analysis and modeling were performed with the R 3.4.3 statistical language environment [20] using of the R-packages: adehabitatHS [21] for ENFA computation and dismo [22,23] for MaxEnt.

Results:

Culicoides occurrences were often positively correlated to precipitation variables and livestock densities; and negatively correlated to the digital elevation model and temperature indices (Fig. 2). Dakar region, the Niayes area and the basin arachidier (groundnut basin) remain the most suitable habitats predicted (Fig. 3). The digital elevation model, the maximum temperature of warmest month, the precipitation of warmest quarter, the mean temperature of wettest quarter, the temperature seasonality, the precipitation of wettest quarter and the goat density were among the most important variables to predict suitable habitats of culicoides (Fig. 4). The AUC values were closed to 0.80 and all models had high accuracy with low p-value (Table 2; Fig. 5).

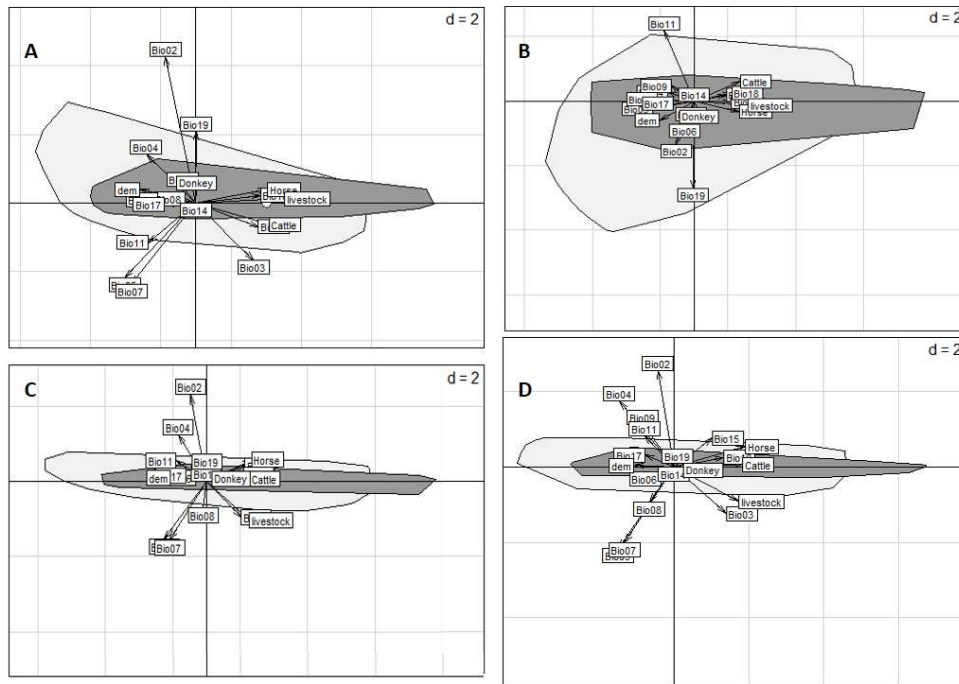


Figure 2: Ecological niche factor analysis (ENFA) of culicoides distribution in Senegal. *C. imicola* (A), *C. oxysoma* (B), *C. enderleini* (C) and *C. miombo* (D). The light gray polygon shows the overall environmental conditions available in the study area, the dark gray one shows environmental conditions where culicoides were observed (representation of the realized niche), and the small white circle corresponds to the barycenter of its distribution.

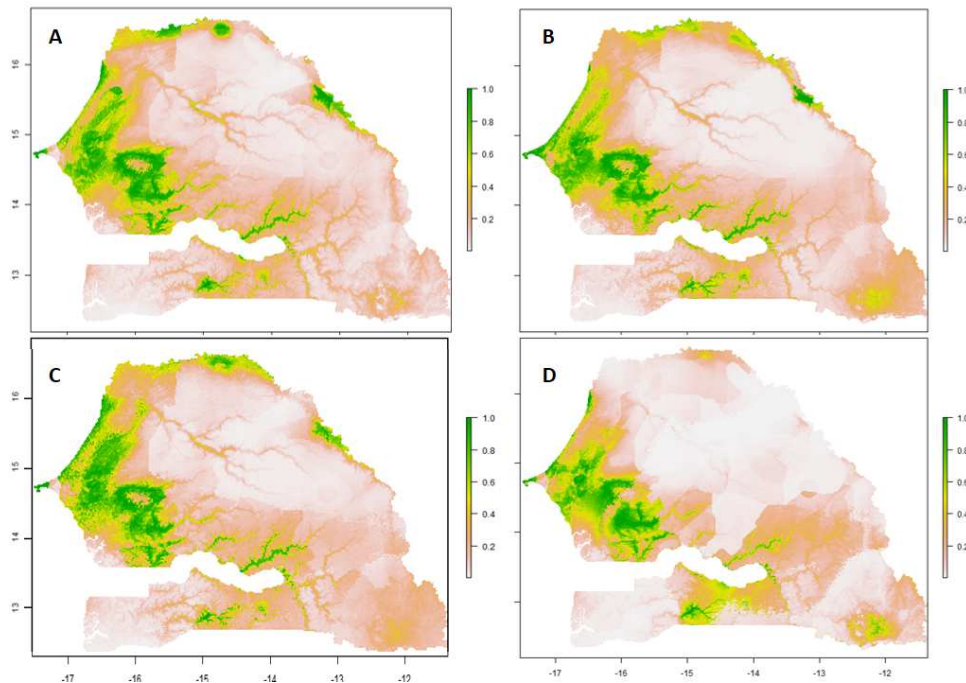


Figure 3: Maxent predicted suitable areas. *C. imicola* (A), *C. oxysoma* (B), *C. enderleini* (C) and *C. miombo* (D). Green areas indicate areas that are likely to have suitable habitat for this vector species while lighter areas indicate areas of that are less suitable for the vector.

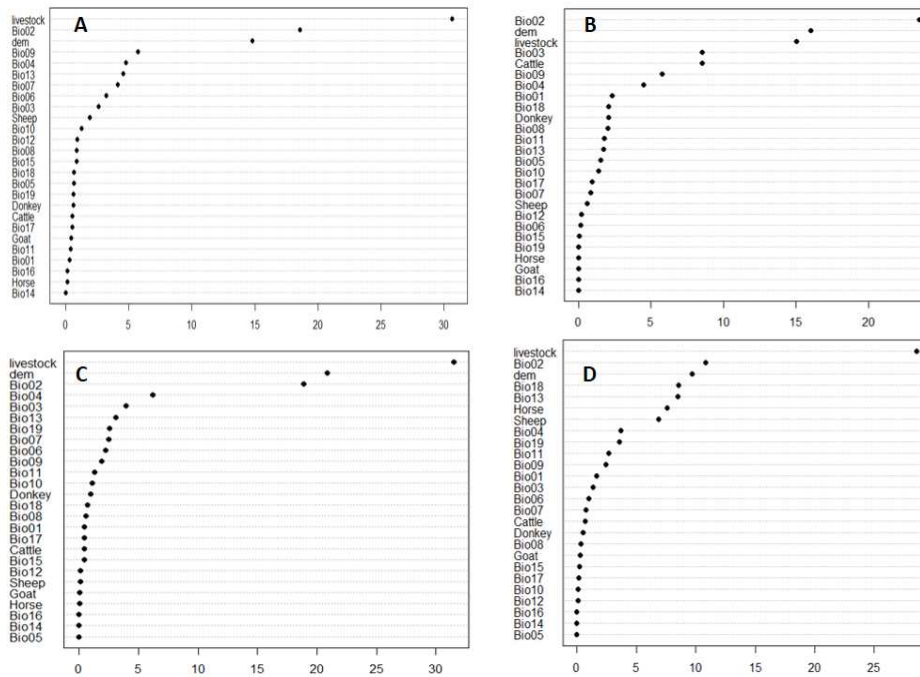


Figure 4: Contribution (%) of each variable for the building of the Maxent models. *C. imicola* (A), *C. oxystoma* (B), *C. enderleini* (C) and *C. miombo* (D).

Table 2: Accuracy of the Niche Models. Test points were used to evaluate omission and 10,000 background points were used to evaluate omission. The reported test omission rate is for equal sensitivity and specificity. The percent contribution of each variable to the models reflects the increase in regularized gain when added to the contribution of the corresponding variable.

Species	Training AUC	Test AUC	Tr.OR	Te.OR	p-value
<i>C. imicola</i>	0.899	0.763 ± 0.042	0.000	0.000	1.393 e ⁻²
<i>C. oxystoma</i>	0.902	0.799 ± 0.039	0.000	0.037	3.194e ⁻²
<i>C. enderleini</i>	0.890	0.772 ± 0.039	0.000	0.000	1.349 e ⁻²
<i>C. miombo</i>	0.921	0.741 ± 0.047	0.000	0.043	5.257 e ⁻³

Abbreviations: AUC Area Under the Curve, Tr.OR Training omission rate, Te.OR Test omission rate

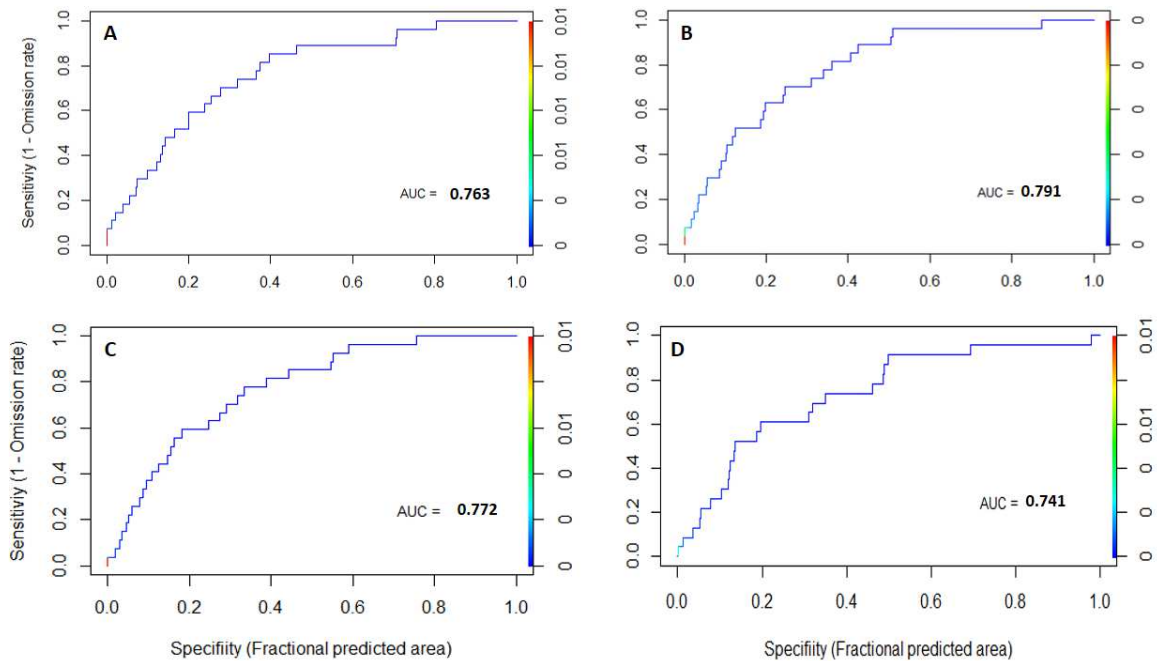


Figure 5: ROC curves and AUCs of the Maxent models. *C. imicola* (A), *C. oxystoma* (B), *C. enderleini* (C) and *C. miombo* (D).

Conclusion:

We presented Senegal scale ecological niche models for *C. imicola*, *C. oxystoma*, *C. enderleini* and *C. miombo* which are susceptible vectors of Bluetongue virus. Our modelling approach was able to determine effects of bioclimatic variables and to generate occurrences of culicoides as risk maps of disease outbreaks.

Keywords:

Vector-borne diseases, Afrotropical region, culicoides, bluetongue, MaxEnt program, ecological niche factor analysis, suitable habitats.

Acknowledgments

The authors are grateful to all the people who gave assistance in operating traps on several nights around the trapping sites. This study was partially funded by the EU grant FP7-261504 EDENext (<http://www.edenext.eu>) and data analysis was funded by Pale-Blu project. The contents of this publication are the sole responsibility of the authors and don't necessarily reflect the views of the European Commission.

References

1. Akakpo A, Wombou Toukam C, Mankor A, Ly C (2011) Economic impact of african horse sickness outbreak in senegal in 2007. Inter-african bureau for animal resources bureau interafrican des ressources animales po Box, NAIROBI, KENYA 59: 1-16.
2. Bakhoum MT (2017) Ecologie et taxonomie intégrative des mouchérons piqueurs du genre# *Culicoides*# Latreille (Diptera: Ceratopogonidae) en région Afrotropicale.
3. Diouf N, Etter E, Lo M, Lo M, Akakpo A (2013) Outbreaks of African horse sickness in Senegal, and methods of control of the 2007 epidemic. *Veterinary Record* 172: 152-152.
4. Diarra M, Fall M, Fall AG, Diop A, Seck MT, et al. (2014) Seasonal dynamics of *Culicoides* (Diptera: Ceratopogonidae) biting midges, potential vectors of African horse sickness and bluetongue viruses in the Niayes area of Senegal. *Parasites & vectors* 7: 147.
5. Boorman J (1989) *Culicoides* (Diptera: Ceratopogonidae) of the Arabian peninsula with notes on their medical and veterinary importance. *Fauna of Saudi Arabia* 10: 160-224.
6. Cornet M, Brunhes J (1994) Révision des espèces de *Culicoides* apparentées à *C. schultzei* (Enderlein, 1908) dans la région afrotropicale (Diptera, Ceratopogonidae). *Bull Soc Entomol Fr* 99: 64.
7. Cornet M, Chateau R (1970) The *Culicoides* of West Africa (second note). Species related to *C. similis* Carter, Ingram & Macfie, 1920 (Diptera, Ceratopogonidae). *Cahiers ORSTOM, Serie Entomologie Medicale at Parasitologie* 8: 141-173.
8. Cornet M, Nevill E, Walker A (1974) Note sur les *Culicoides* (Diptera, Ceratopogonidae) du groupe de *C. milnei* Austen, 1909, en Afrique orientale et australe. *Cah ORSTOM Ser Ent Med Parasitol Off Rech Sci Tech Outre Mer*.
9. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology* 25: 1965-1978.
10. Bakhoum MT, Fall M, Seck M, Gardes L, Fall A, et al. (2016) Foraging range of arthropods with veterinary interest: new insights for Afrotropical *Culicoides* biting midges (Diptera: Ceratopogonidae) using the ring method. *Acta tropica* 157: 59-67.

11. Fall M, Fall AG, Seck MT, Bouyer J, Diarra M, et al. (2015) Host preferences and circadian rhythm of *Culicoides* (Diptera: Ceratopogonidae), vectors of African horse sickness and bluetongue viruses in Senegal. *Acta tropica* 149: 239-245.
12. Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological niche factor analysis: how to compute habitat suitability maps without absence data? *Ecology* 83: 2027-2036.
13. Basille M, Calenge C, Marboutin E, Andersen R, Gaillard J-M (2008) Assessing habitat selection using multivariate statistics: Some refinements of the ecological-niche factor analysis. *Ecological modelling* 211: 233-240.
14. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, et al. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and distributions* 17: 43-57.
15. Merow C, Smith MJ, Silander Jr JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36: 1058-1069.
16. Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161-175.
17. Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of biogeography* 34: 102-117.
18. Young N, Carter L, Evangelista P (2011) A MaxEnt model v3. 3.3 e tutorial (ArcGIS v10). Fort Collins, Colorado.
19. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29-36.
20. R Core Team (2017) R : A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
21. Calenge C (2006) The package "adehabitat" for the R software: a tool for the analysis of space and habitat use by animals. *Ecological modelling* 197: 516-519.
22. Hijmans RJ, Elith J (2013) Species distribution modeling with R. R package version 08-11.
23. Hijmans RJ, Phillips S, Leathwick J, Elith J (2013) dismo: Species distribution modeling. R package version 0.9-3. Available at: <http://CRAN.R-project.org/package=dismo>.