



## Research paper

## Using genetic data to improve species distribution models

Jérémy Bouyer<sup>a,b,\*</sup>, Renaud Lancelot<sup>a,1</sup><sup>a</sup> Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de "Animal, Santé, Territoires, Risques et Ecosystèmes", Campus International de Baillarguet, 34398 Montpellier, France<sup>b</sup> Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Unité Mixte de Recherche 'Interactions hôtes-vecteurs-parasites-environnement dans les maladies tropicales négligées dues aux trypanosomatides', 34398 Montpellier, France

## ARTICLE INFO

## Article history:

Received 8 October 2016

Received in revised form 21 March 2017

Accepted 22 March 2017

Available online 23 March 2017

## Keywords:

Friction

Movement

Population genetics

Landscape genetics

Maxent

## ABSTRACT

Tsetse flies (Diptera, Glossinidae) transmit human and animal trypanosomoses in Africa, respectively a neglected human disease (sleeping sickness) and the most important constraint to cattle production in infested countries (*nagana*). We recently developed a methodology to map landscape friction (i.e. resistance to movement) for tsetse in West Africa. The goal was to identify natural barriers to tsetse dispersal, and potentially isolated tsetse populations for targeting elimination programmes. Most species distribution models neglect landscape functional connectivity whereas environmental factors affecting suitability or abundance are not necessarily the same as those influencing gene flows. Geographic distributions of a given species can be seen as the intersection between biotic (B), abiotic (A) and movement (M) factors (BAM diagram).

Here we show that the suitable habitat for *Glossina palpalis gambiensis* as modelled by Maxent can be corrected by landscape functional connectivity (M) extracted from our friction analysis. This procedure did not degrade the specificity of the distribution model ( $P = 0.751$ ) whereas the predicted distribution area was reduced. The added value of this approach is that it reveals unconnected habitat patches.

The approach we developed on tsetse to inform landscape connectivity (M) is reproducible and does not rely on expert knowledge. It can be applied to any species: we call for a generalization of the use of M to improve distribution models.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The geographic distribution of a given species can be seen as the intersection between biotic (B), abiotic (A) and movement (M) factors (BAM diagram (Soberón and Peterson, 2005)). However, most species distribution models neglect the latter (Barve et al., 2011), often because they are based on presence/absence data and rarely include movements components. We recently developed a methodology allowing landscape friction (i.e. resistance to movement) for tsetse to be mapped (Bouyer et al., 2015b). Here we want to show that it is important to correct A by M, which is itself linked to B (density dependent dispersal, competition, etc.).

Tsetse flies (Diptera, Glossinidae) transmit human and animal trypanosomoses in Africa, respectively causing a neglected human disease (sleeping sickness) and the most important constraint to cattle production in infested countries (*nagana*): they are the target of the

Pan African Tsetse and Trypanosomoses Eradication Campaign (PATTEC). However, tsetse eradication efforts were only sustainable in the past when targeting isolated populations (Vreysen et al., 2013b). Recently, we used genetic distances between 37 western African tsetse populations of *Glossina palpalis gambiensis* (Gpg) and remotely-sensed environmental data to map landscape friction, which is related to the functional connectivity of the habitat (i.e. how far animal species can move in the environment among patches of suitable habitat). Friction can be used to quantify a large set of movements parameters including distance, travel time, routes trade, etc. In population genetics, friction is modelled to (i) identify landscape and environmental features constraining genetic connectivity, (ii) elucidate the ecological processes influencing spatial genetic structure, and (iii) assess genetic connectivity and predict how future landscape changes might influence it (Spear et al., 2010). In our case, we used it to identify natural barriers to tsetse dispersal, and potentially isolated tsetse populations for targeting elimination programs (Bouyer et al., 2015b). Landscape functional connectivity involves rather more components than friction, particularly the distance between patches. However, provided that the spatial resolution of predictions accounts for the species ecology i.e. the prediction pixels are not too small to represent barriers (see below), friction can be used as a good proxy of connectivity as in the present study.

\* Corresponding author at: Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de "Animal, Santé, Territoires, Risques et Ecosystèmes", Campus International de Baillarguet, 34398 Montpellier, France.

E-mail address: [bouyer@cirad.fr](mailto:bouyer@cirad.fr) (J. Bouyer).

<sup>1</sup> Equal contributions by authors.

Models of habitat suitability or abundance may not be adequate proxies for gene flows (Peterman et al., 2014). In our tsetse study, we independently modelled habitat suitability and friction (landscape resistance to tsetse movements), the former based on 8,941 presence-absence data, and the latter derived from the genetic distance between 37 tsetse populations. We observed that environmental factors driving these habitat properties were different (Bouyer et al., 2015b): on the one hand, the average maximum land surface temperature and annual rainfall were the main drivers of habitat suitability; on the other hand, friction predictors were landscape fragmentation metrics and the presence of a watershed. Thus, the spatial distribution of this tsetse species in western Africa was compatible with Hanski-Ovaskainen's metapopulation model (Hanski and Ovaskainen, 2000). The two parameters of this model are the extinction rate - which can be fitted by a population dynamic model depending on habitat suitability, and the colonization rate - which is related to patch areas as well as the Euclidian distance and habitat connectivity between patches (Peck, 2012). This observation has important implications for tsetse control because the extinction probability associated with control measures might be altered when taking into account the spatial structure of the tsetse metapopulation (Peck, 2012). In Senegal, we observed that some habitat patches predicted as suitable for Gpg by the Maxent model were not inhabited by tsetse because they were previously targeted by eradication efforts, and separated from the main tsetse population by a few kilometres of unsuitable habitat preventing their reinvasion (Fig. 1) (Dicko et al., 2014). Therefore, landscape friction might explain why unconnected patches of suitable habitat either could not be colonized by the species, or did not benefit from a high enough colonization rate to counteract the extinction rate.

**2. Material and methods**

To compute M, we used the landscape friction map F described above (Bouyer et al., 2015b). The spatial resolution of this map was set to ~5 km because smaller bands of high friction would not be enough to act as barriers to fly movement. Maxent predictions (A layer) were thus rescaled at the same spatial resolution. To do so, we selected the

habitat suitability threshold (number of suitable pixels of ~1 km within pixels of ~5 km) to keep a sensitivity of 96% in the aggregated layer with respect to the initial resolution. We identified a friction threshold  $\tau$ , defining M as  $F \leq \tau$ . This threshold was selected by maximizing the specificity for a given sensitivity read from receiver operating characteristic (ROC) curve. This estimate was fine-tuned at the maximum value for which known marginal tsetse populations (in coastal Senegal and Bijagos islands, Bissau Guinea) were still considered as living in suitable habitat with low friction. This information was confirmed previously based on a systematic entomological sampling survey conducted within the feasibility phase of a tsetse eradication program (Bouyer et al., 2010).

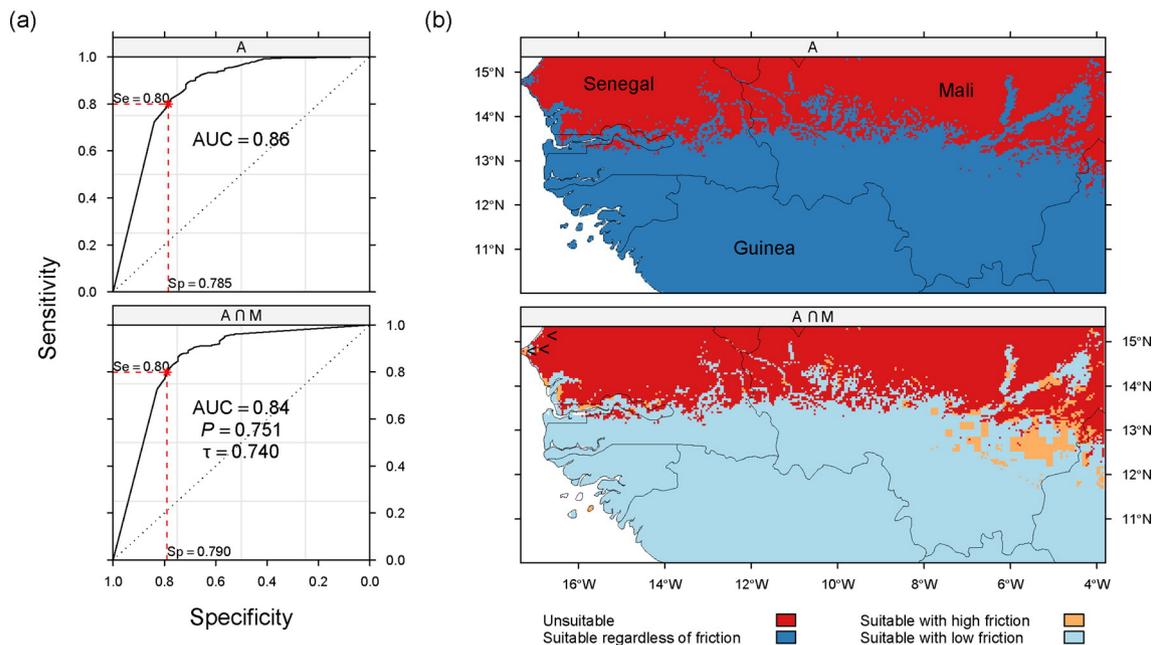
**3. Results**

Fig. 1A presents the area under the ROC curve (AUC) of two different models, the first computed from a presence-absence dataset only (A layer) and the second corresponding to the  $A \cap M$  layer. The AUC values obtained for the two models are similar and for an 80% sensitivity, this procedure did not degrade the specificity of the distribution model ( $P=0.751$ ) whereas the overall predicted distribution area was reduced.

In Fig. 1B, we show the suitable habitat for Gpg predicted by the Maxent model (upper map), as well as the intersection of suitability (A) and friction (M) that highlights unsuitable habitat, suitable habitat surrounded by areas of high friction for tsetse genetic flows, and connected suitable habitat (lower map). Habitat classed as "suitable with high friction" can be considered as tsetse-free despite suitable abiotic conditions because tsetse are unable to disperse between suitable patches. This result was corroborated in the Niayes area of Senegal, where the habitat patches indicated by a "<" in Fig. 1B are suitable but not inhabited anymore by tsetse.

**4. Discussion**

Here we corrected a tsetse distribution model using a friction map developed to explain genetic distances between Gpg populations. The



**Fig. 1.** Distribution of *Glossina palpalis gambiensi* in West Africa. (a) Receiver operating characteristic (ROC) curves of Maxent model (A layer), and the Maxent model corrected by a friction model ( $A \cap M$  layer, see text for details). (b) Predictions of suitable habitat patches by the Maxent model at a spatial resolution of ~5 km on the top panel, corrected by the friction model on the bottom. Some habitat patches indicated by a "<" symbol are suitable but located in high friction areas, and are not inhabited anymore by tsetse. Abbreviations: A, abiotic; M, movement; AUC, area under the ROC curve; P, probability that the difference in AUC of the two models was observed by chance;  $\tau$ , friction threshold defining M; Se, sensitivity; Sp, specificity.

added value of this approach is best seen in the eastern part of the map where unconnected habitat patches are revealed on the northern edge of the tsetse belt. These habitat patches might be the target of forthcoming Gpg eradication programmes (Vreysen et al., 2013b). Such programs were demonstrated to be cost-effective on isolated islands like Zanzibar against *Glossina pallidipes* (Vreysen et al., 2014) or in continental ecological islands like the Niayes area of Senegal (at the top left of tsetse distribution in Fig. 1B) against *G. p. gambiensis* (Bouyer et al., 2015a; Bouyer et al., 2014). Unfortunately, such sites are representing less than 5% of our study area. In the remaining area, integrated control of trypanosomoses will probably be a better option (Bouyer et al., 2013).

Interestingly, the sites corresponding to “suitable with high friction” in the Niayes areas were previously infested by *G. p. gambiensis* in the 1970s. This area was then subjected to two control programs using residual spraying of dieldrin and insecticide targets (Morel and Touré, 1967; Touré, 1973, 1974). Although the former control programs failed to eradicate the full target population, they probably eliminated some pockets of *G. p. gambiensis* that were completely isolated from the other infested areas by urban settlements or sand dunes. The absence of *G. p. gambiensis* in these sites was confirmed during intensive sampling for several months and using numerous traps within the feasibility study of the new eradication project that started in 2007 (Bouyer et al., 2010), despite these sites appearing fully suitable for this species based on phytosociological criteria (Bouyer et al., 2005). In Mali, using a mark-release-recapture experiment we confirmed that tsetse dispersal into unsuitable landscapes is actually very low for both male and female *G. p. gambiensis* (Vreysen et al., 2013a). To our knowledge, this is a rare case of documentation of the metapopulation theory (Hanski and Ovaskainen, 2000).

The approach we developed on tsetse to quantify landscape friction (M) is reproducible and does not rely on expert knowledge (Bouyer et al., 2015b). It can be applied to any species, not only to improve control of vector and pest species, but also to improve the protection of endangered species, on a planet with ever increasing fragmentation of landscapes: we call for a generalization of the use of M to improve distribution models.

## Acknowledgements

This work was supported by the research project “Integrated Vector Management: innovating to improve control and reduce environmental impacts” of Institut Carnot Santé Animale excellence network.

## References

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J., Villalobos, F., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Model.* 222, 1810–1819.

- Bouyer, J., Bouyer, F., Donadeu, M., Rowan, T., Napier, G., 2013. Community- and farmer-based management of animal African trypanosomosis in cattle. *Trends Parasitol.* 29, 519–522.
- Bouyer, F., Bouyer, J., Seck, M.T., Sall, B., Dicko, A.H., Lancelot, R., Chia, E., 2015a. Importance of vector-borne infections in different production systems: bovine trypanosomosis and the innovation dynamics of livestock producers in Senegal. *Revue scientifique et technique (International Office of Epizootics)*. 34, pp. 199–212.
- Bouyer, J., Dicko, A.H., Cecchi, G., Ravel, S., Guerrini, L., Solano, P., Vreysen, M.J.B., De Meeüs, T., Lancelot, R., 2015b. Mapping landscape friction to locate isolated tsetse populations candidate for elimination. *Proc. Natl. Acad. Sci. U. S. A.* 112, 14575–14580.
- Bouyer, J., Guerrini, L., César, J., de la Rocque, S., Cuisance, D., 2005. A phyto-sociological analysis of the distribution of riverine tsetse flies in Burkina Faso. *Med. Vet. Entomol.* 19, 372–378.
- Bouyer, F., Seck, M.T., Dicko, A., Sall, B., Lo, M., Vreysen, M., Chia, E., Bouyer, J., Wane, A., 2014. Ex-ante cost-benefit analysis of tsetse eradication in the Niayes area of Senegal. *PLoS Negl. Trop. Dis.* 8, e3112.
- Bouyer, J., Seck, M.T., Sall, B., Guerrini, L., Vreysen, M.J.B., 2010. Stratified entomological sampling in preparation of an area-wide integrated pest management programme: the example of *Glossina palpalis gambiensis* in the Niayes of Senegal. *J. Med. Entomol.* 47 (4), 543–552.
- Dicko, A.H., Lancelot, R., Seck, M.T., Guerrini, L., Sall, B., Lo, M., Vreysen, M.J.B., Lefrançois, T., Williams, F., Peck, S.L., Bouyer, J., 2014. Using species distribution models to optimize vector control: the tsetse eradication campaign in Senegal. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10149–10154.
- Hanski, I., Ovaskainen, O., 2000. The metapopulation capacity of a fragmented landscape. *Nature* 404, 755–758.
- Morel, P.C., Touré, S., 1967. *Glossina palpalis gambiensis* Vanderplank 1949 (Diptera) dans la région des Niayes et sur la Petite Côte (République du Sénégal). *Rev. Elev. Méd. vét. Pays trop.* 20, 571–578.
- Peck, S.L., 2012. Networks of habitat patches in tsetse fly control: implications of metapopulation structure on assessing local extinction. *Ecol. Model.* 246, 99–102.
- Peterman, W.E., Connette, G.M., Semlitsch, R.D., Eggert, L.S., 2014. Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Mol. Ecol.* 23, 2402–2413.
- Soberón, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Inform.* 2, 1–10.
- Spear, S.F., Balkenhol, N., Fortin, M.-J., McRae, B.H., Scribner, K., 2010. Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Mol. Ecol.* 19, 3576–3591.
- Touré, S., 1973. Lutte contre *Glossina palpalis gambiensis* dans la région des niayes du Sénégal. *Rev. Elev. Méd. vét. Pays trop.* 26, 339–347.
- Touré, S., 1974. Note sur quelques particularités dans l'habitat de *Glossina palpalis gambiensis* Vanderplank, 1949 (Diptera, Glossinidae) observées au Sénégal. *Rev. Elev. Méd. vét. Pays trop.* 27, 81–94.
- Vreysen, M., Balenghien, T., Saleh, K.M., Maiga, S., Koudougou, Z., Cecchi, G., Bouyer, J., 2013a. Release-recapture studies confirm dispersal of *Glossina palpalis gambiensis* between river basins in Mali. *PLoS Negl. Trop. Dis.* 7, e2022.
- Vreysen, M.J.B., Saleh, K., Mramba, F., Parker, A., Feldmann, U., Dyck, V.A., Msangi, A., Bouyer, J., 2014. Sterile insects to enhance agricultural development: the case of sustainable tsetse eradication on Unguja Island, Zanzibar using an area-wide integrated pest management approach. *PLoS Negl. Trop. Dis.* 8, e2857.
- Vreysen, M.J.B., Seck, M.T., Sall, B., Bouyer, J., 2013b. Tsetse flies: their biology and control using area-wide integrated pest management approaches. *J. Invertebr. Pathol.* 112, S15–S25.