Section 2
Tsetse distribution in the Mouhoun river basin (Burkina Faso): the role of global and local geospatial datasets

Laure Guerrini[1,2*], Issa Sidibé[2] and Jérémy Bouyer[2,3]

[1]Centre de coopération internationale en recherche agronomique pour le développement, Département Es

[2]Centre International de Recherche-Développement sur l'Elevage en zone Subhumide

[3]Centre de coopération internationale en recherche agronomique pour le développement Département BIOS

[*]Corresponding author: L. Guerrini. E-mail: laure.guerrini@fasonet.bf

ABSTRACT
In this study, we explore the potential of using select global geographic information system (GIS) datasets to map the distribution and densities of riverine tsetse fly at a local scale in the Mouhoun river basin (Burkina Faso). In particular, we analyse the correlation between low-resolution datasets that predict global and regional tsetse distribution and more than 800 trapping scores for *Glossina palpalis gambiensis* Vanderplank and *G. tachinoides* Westwood. The results show that these datasets with global or regional scope, including the Global Land Cover database for the year 2000 (GLC2000), are not suitable for use at a local scale (e.g. in designing baseline data collection protocols for vector control campaigns). On the other hand, higher resolution global datasets available in the public domain, namely the NASA Landsat Orthorectified Image Library (LOIL) and the HydroSHEDS drainage network, have great potential for mapping tsetse habitat when used in process-based models.

INTRODUCTION
In Burkina Faso, as in most sub-Saharan West African countries that are inhabited by tsetse flies, animal African trypanosomiasis (AAT) is a major hindrance to cattle breeding (Itard, Cuisance and Tacher, 2003). Tsetse flies are also cyclic vectors of sleeping sickness in humans, spreading human African trypanosomiasis (HAT). Two riverine tsetse species, *Glossina palpalis gambiensis* Vanderplank 1949 (Diptera, Glossinidae) and *G. tachinoides* Westwood 1850, are still present in considerable densities.

The link between environment and the presence or abundance of vectors of trypanosomiasis is well known, and the use of remote sensing has become an essential tool for the epidemiological analysis of vector-borne diseases (Rogers and Randolph, 1993; Rogers, Hay and Packer, 1996; Robinson, Rogers and Brian, 1997; Hendrickx *et al.*, 1999; de La Rocque *et al.*, 2005). In Burkina Faso, recent studies have demonstrated the relationship between the riverine forest ecotype (and its disturbance level) and the abundance of riverine tsetse. These studies follow a theory regarding the riverine forest ecotype (Morel, 1983) that was later enhanced by the integration of human-driven disturbance (Bouyer *et al.*, 2005).

In this paper, we examine the usefulness of various global and local geospatial datasets to map the habitat, distribution and densities of riverine tsetse, with the ultimate goal of estimating AAT risk at a river basin scale. The types of datasets examined are:

- tsetse distribution (PAAT-IS);
- land cover (GLC2000);
- medium-resolution satellite imagery (Landsat 7 ETM+);
- hydrographic network (HydroSHEDS);
- rainfall (FAOClim).

TSETSE DISTRIBUTION MAPS OF THE PAAT-IS: A COMPARISON WITH FIELD DATA ON A LOCAL SCALE

A recent tsetse elimination initiative has been launched in the Mouhoun river basin within the framework of the Pan African Tsetse and Trypanosomiasis Eradication Campaign (PATTEC), with the financial support of the African Development Bank. The maps that are available in the Programme Against African Trypanosomiasis Information System (PAAT-IS)[1] (see page 3) were used to design the initial area in which baseline data collection would occur. For the Mouhoun river basin, PAAT-IS maps are available at 5 km and 1 km resolutions. These maps consist of probabilities of the presence of different tsetse species. In our case study, the species of interest are *G. p. gambiensis* and *G. tachinoides*. The PAAT-IS maps have been used widely to support strategic decision-making at global and regional levels (e.g. for the selection of priority areas for intervention). However, their potential and limitations for interventions at a local scale have not been evaluated comprehensively.

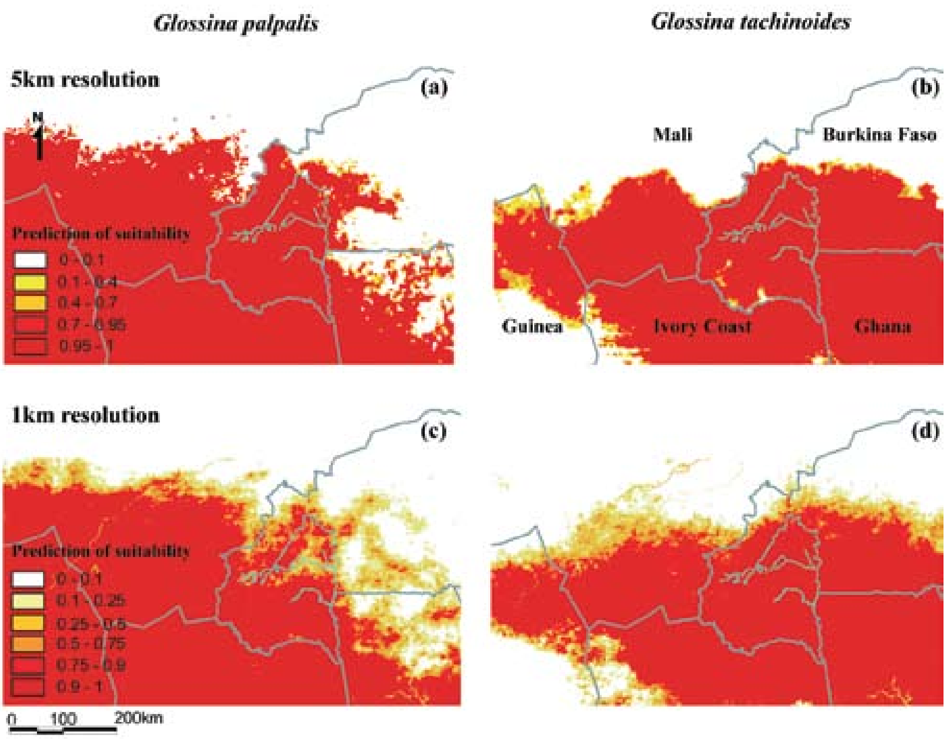
We compared the predictions of the PAAT-IS models to field data collected during various research projects led by the Centre International de Recherche-Développement en zone Subhumide (CIRDES). The most common way to interpret PAAT-IS maps is to assume that tsetse are present in a given area when the probability of suitability is 0.5 (i.e. 50 percent) or greater and that they are absent when the probability is less than 0.5.

Figure 1 shows at resolutions of 5 km and 1 km the predicted areas of suitability for *G. palpalis* and *G. tachinoides*. (It must be noted that the PAAT-IS maps at 5 km resolution, which are available at a continental level, treat the three *G. palpalis* subspecies as a single species.) Figure 2 presents the percentages of trapping sites at which tsetse were actually caught, grouped by class of probability (as derived from the PAAT-IS 5 km and 1 km resolution models, respectively).

The correlation between field data and the 5 km resolution models is poor. In the case of *G. p. gambiensis*, tsetse were found at 32 trapping sites where the probability of suitability, as estimated by the models, was 0.1. Similarly, no tsetse were caught at 3 trapping sites where the estimated probability of suitabilities ranged from 0.6 to 0.9. In the case of *G. tachinoides*, no flies were caught at any sites with a probability of suitability less than 1. If not adequately interpreted, the use of such predictions for

[1] http://www.fao.org/ag/againfo/programmes/en/paat/maps.html

FIGURE 1
Predicted areas of suitability for *G. palpalis* and *G. tachinoides* at map resolutions of 5 km (a and b, respectively) and 1 km (c and d, respectively). Source: PAAT-IS

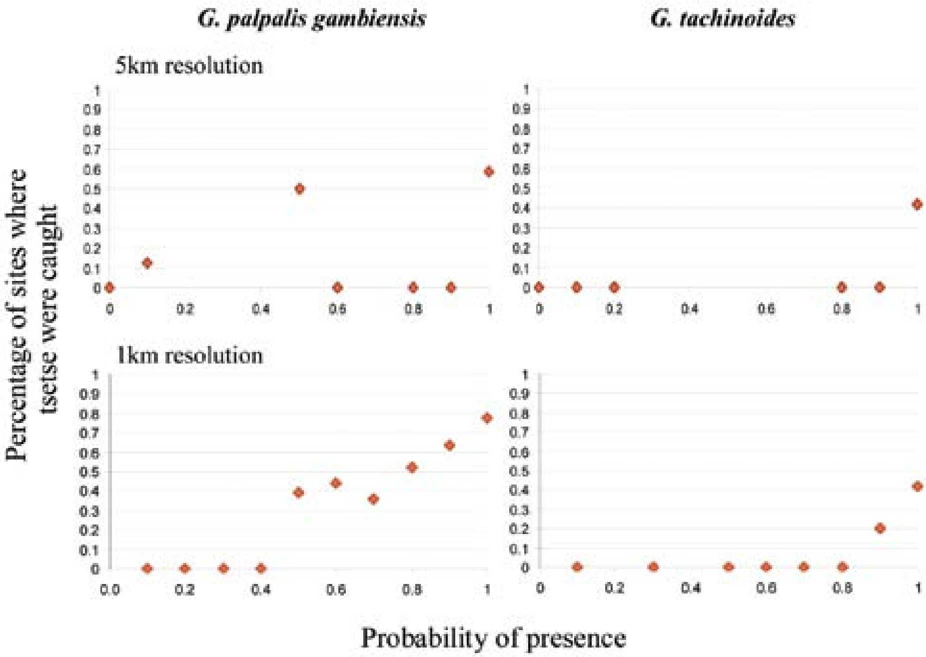


planning the deployment of tsetse traps for baseline data collection would lead to an important overestimation of the sampling area, resulting in significant economic losses (i.e. the funds wasted on misplaced traps).

The predictions at 1 km resolution provide substantially better results than the predictions at 5 km resolution. If a threshold of 0.5 is used to distinguish between absence and presence, the predictions for *G. p. gambiensis* are better than those for *G. tachinoides*. However, if a different threshold is used (e.g. 0.8), the PAAT-IS map for *G. tachinoides* may also prove to be useful.

Even though this case study utilizes field data obtained over a relatively limited geographical area, especially if compared with the global and regional scope of the PAAT-IS maps, it nevertheless demonstrates the limits of PAAT-IS predictions in supporting decision-making at the local level. We argue that when moving on from strategic decision-making towards planning and implementation of field activities at a local level, low-resolution geospatial datasets are inadequate and should only be used with extreme caution. Clearly, the need exists for datasets of higher resolution and for novel analytical methods to map the distribution of tsetse flies.

FIGURE 2
Percentage of 764 trapping sites at which tsetse were caught, grouped by class of probability (as derived from the PAAT-IS 5 km and 1 km resolution models, respectively)



GLOBAL LAND COVER 2000

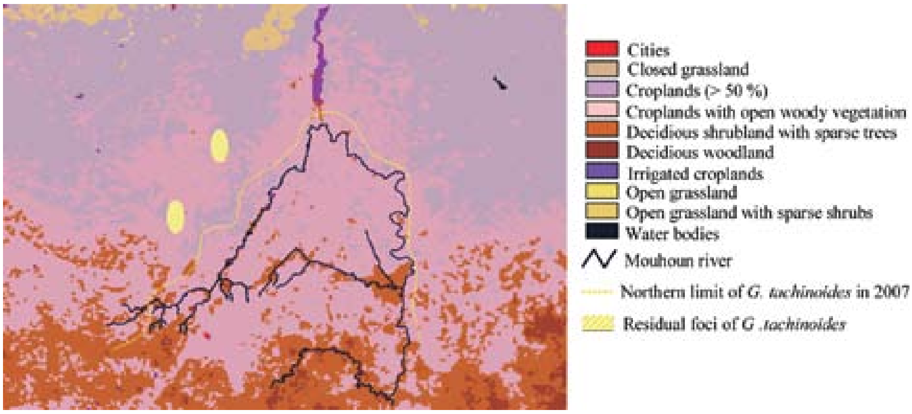
Land cover is arguably the most relevant environmental parameter affecting the suitability of habitat for tsetse flies in that vegetation is either directly or indirectly shaped by soils, climate and human activities. Land cover is also one of the indicators of human intervention on the land most easily detected by remote sensing.

Before the release of the 300 m resolution GlobCover[2] land cover map of the world on 30 September 2008, the 1 km resolution GLC2000[3] provided the most detailed picture of land cover of the earth (see page 9). The GLC2000 land cover classes were found to be statistically correlated with the continental distribution of the three tsetse fly subgenera (*fusca*, *palpalis* and *morsitans*) (Cecchi *et al.*, 2008). In particular, 56 percent of the distribution of the *fusca* group and 46 percent of the distribution of the *palpalis* group could be predicted by GLC2000 land cover classes.

We analysed the potential of the GLC2000 to predict the local presence and abundance of *G. p. gambiensis* and *G. tachinoides* (both belonging to the *palpalis* group) in the Mouhoun river basin. As an example, Figure 3 presents the northern limit of the distribution of *G. tachinoides* superimposed on the GLC2000 land cover units. Four

[2] http://ionia1.esrin.esa.int/index.asp
[3] http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php

FIGURE 3
Northern border of the distribution area of G. tachinoides (in yellow) superimposed on GLC2000 vegetation units



units are present within the studied area: "croplands with open woody vegetation", "deciduous shrubland with sparse trees", "croplands (> 50 percent)" and "irrigated croplands". Table 1 shows the mean apparent densities per trap per day (ADT) for both species calculated for the four land cover units, together with the probability of species presence derived from fly catches at 831 trapping sites.

These results are in line with the continental level study, which assigned to the GLC2000 classes in Table 1 a degree of suitability for the *palpalis* group that ranged from low to moderate. Somewhat unexpectedly, the maximum ADTs for both species are observed in the more disturbed units corresponding to "croplands (> 50 percent)". On the other hand, the maximum probabilities of presence are found in classes with a higher proportion of natural vegetation, namely "deciduous shrublands with sparse trees" for *G. p. gambiensis* and "croplands with open woody vegetation" for *G. tachinoides*. Considering that the presence and abundance of tsetse are known to be negatively correlated with the disturbance of riverine forests caused by agriculture (Bouyer *et al.*, 2005), these results call for an explanation.

In Africa, the discrimination of agricultural areas from natural vegetation using satellite imagery of 1 km resolution is quite problematic because of the characteristics of prevailing farming systems and the spatial pattern of croplands. As a result, GLC2000 resolution is probably inadequate to depict the rather small and scattered patches of riverine vegetation that are so important for tsetse fly, especially in this area at the limit of its distribution. This is confirmed by Figure 4, which presents a land cover map at 30 m resolution generated through supervised classification of a Landsat 7 ETM+ scene, as compared with the GLC2000. The forest unit, which is the most important vegetation unit for mapping of tsetse distribution, is completely absent from the GLC2000 and

TABLE 1

Mean apparent density per trap per day (ADT) for *G. p. gambiensis* and *G. tachinoides* (standard deviation in brackets), together with the probability of species presence, in four GLC2000 vegetation units (Mouhoun river basin, Burkina Faso)

GLC2000 units	ADT *G.p.g.*	ADT *G.t.*	Number of traps	Absence of *G.p.g.*	Absence of *G.t.*	Probability of presence of *G.p.g.*	Probability of presence of *G.t.*
Croplands with open woody vegetation	2.79 (5.90)	2.79 (5.27)	574	306	276	0.47	0.52
Deciduous shrubland with sparse trees	2.78 (3.60)	0.56 (1.18)	210	50	144	0.76	0.31
Croplands (> 50%)	6.93 (13.63)	4.26 (14.36)	27	16	16	0.41	0.41
Irrigated croplands	1.80 (2.24)	0.05 (0.22)	20	05	19	0.75	0.05

seems to be have been diluted either in the class "deciduous shrubland with sparse trees" or in "croplands with open woody vegetation".

Our results confirm the limitations of using coarse resolution global datasets at a local scale and the need for an approach using multiple resolutions to study the link between land cover maps and tsetse habitat.

MEDIUM-RESOLUTION SATELLITE IMAGERY (LANDSAT 7 ETM+)

The previous two sections have highlighted some of the limits of using global datasets at a local scale. We summarize here the methodology and results of a recent study to assess riverine tsetse fly densities and trypanosomiasis risk along the Mouhoun river using medium-resolution satellite imagery (Landsat ETM+) coupled with entomological and environmental field data (Bouyer *et al.*, 2006; Guerrini and Bouyer, 2007; Guerrini and Bouyer, in press). The methodology was based on an initial distinction of three ecological sections along the Mouhoun river (Guinean, Sudano-Guinean and Sudanese gallery forest) and the discrimination of peri-riverine landscape units within a 1 km buffer of the river.

The aim of the landscape classification method was to identify clusters corresponding to three entomological landscapes described in the field (protected forest, border of a protected forest, and cultivated and grazed areas) using seven land-use classes obtained from supervised classification of Landsat 7 ETM+ imagery (Figure 4c). Riverine forests are often too thin (< 10 m) to be clearly detected with Landsat 7 multispectral bands (30 m) and were thus analysed through their neighbouring pixels. Points were randomly generated within each ecological section and then analysed to identify clusters of similar neighbourhoods. The areas of each land-use class were calculated in buffers of 500 m around each point and then expressed as a percentage of the total buffer area (Figure 5).

Clusters of similar neighbourhoods were identified from the hierarchical classification and then matched with the entomological landscapes described in the field. This approach

FIGURE 4

(a) GLC2000 vegetation units, (b) false-colour composition (TM4, TM3, TM2) of Landsat 7 ETM+ scene and (c) associated supervised classification of the northern edge of the Mouhoun river loop (Burkina Faso) within a 1 km buffer of the river. Source: Guerrini and Bouyer, 2007

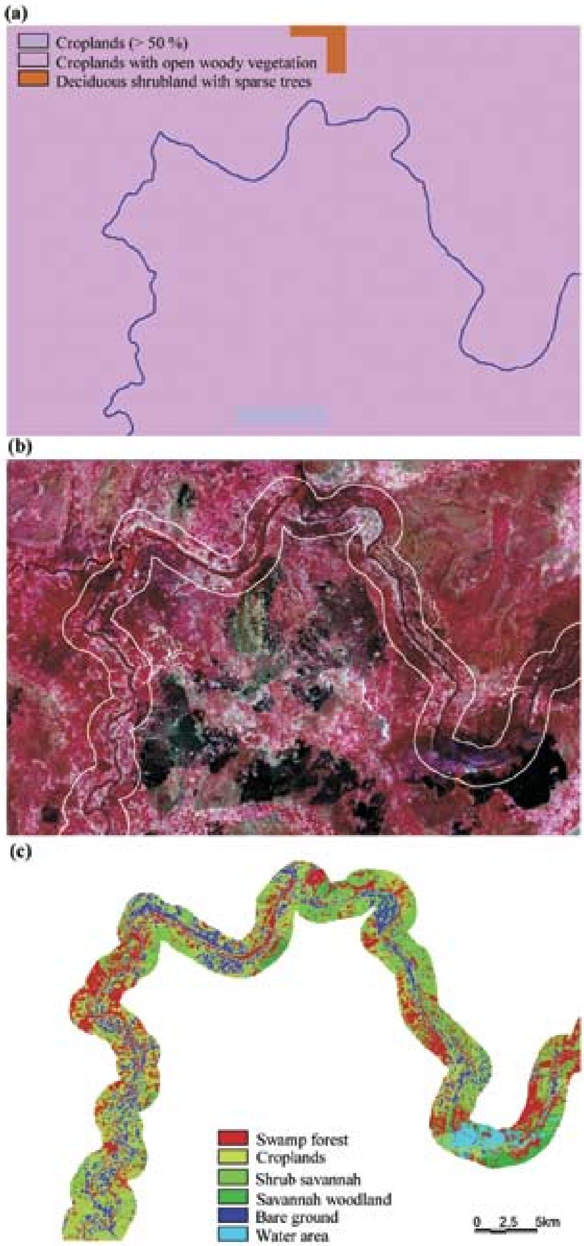
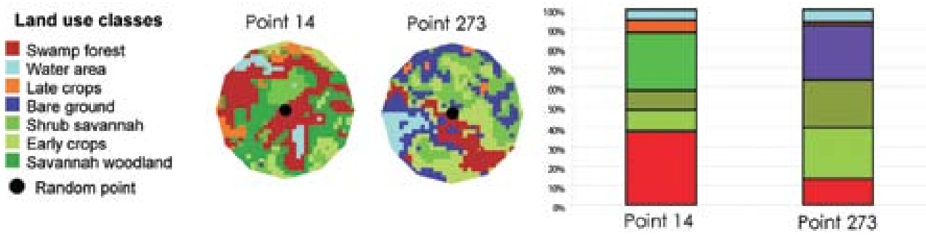


FIGURE 5
Percentages of seven vegetation units derived from supervised classification in 500 m buffers around two points randomly generated along the river course. Source: Guerrini and Bouyer, 2007



allowed classification of the river network into three disturbance levels (disturbed, half-disturbed and natural). The surface of the water around the river course could be used to predict the ecotype of the riverine forest (Guinean, Sudano-Guinean and Sudanese) (Bouyer *et al.*, 2005). Overall, a good classification was obtained for 81 percent of the sites (Guerrini *et al.*, 2008). Figure 6 presents the distribution of these disturbance levels and ecotypes along the Mouhoun river loop.

The ecotype and disturbance levels were matched to obtain homogeneous landscapes in which the ADT of the two species was measured by means of 689 trapping sites (Guerrini and Bouyer, 2007). The ADTs were then mapped along the main course of the Mouhoun river and two of its tributaries (the Leyessa and the Balé), as shown in Figure 6.

Our model predictions were validated against an independent dataset (66 trapping sites). A very good correlation between the model outputs and the field data was observed (Kendall Test results for *G. p. gambiensis*: $T = 0.37$, $z = 4.19$, $p = 2.831e\text{-}05$; Kendall Test results for *G. tachinoides*: $T = 0.39$, $z = 4.67$, $p = 3.036e\text{-}06$).

The method briefly outlined here demonstrates the potential of medium-resolution satellite imagery as a tool in tsetse fly mapping and subsequent assessment of AAT risk at a local level.

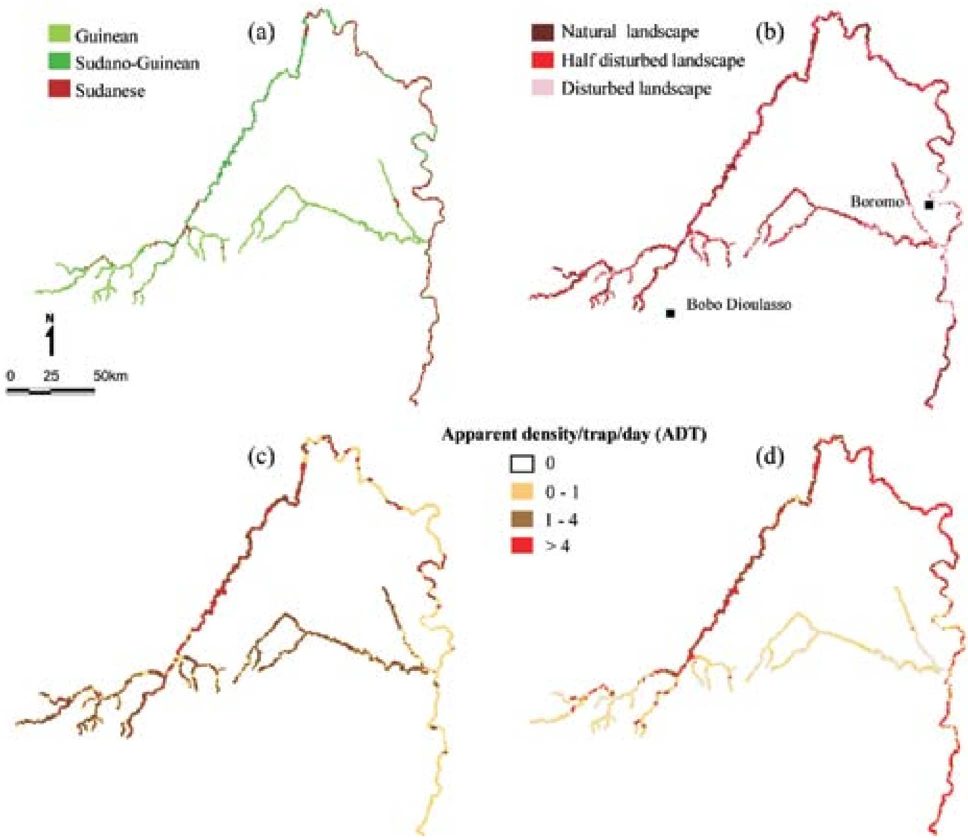
HYDROGRAPHIC NETWORK

A recent study (Guerrini and Bouyer, 2007) compared the HydroSHEDS[4] drainage network (see page 17) with a hydrographic network, digitized from Landsat 7 ETM+ satellite imagery, that was used to map tsetse densities in Burkina Faso. Figure 7 presents the two different sets of vector data.

Although no systematic or quantitative analysis was carried out, visual analysis indicates a very good match between the two datasets. The maximum observed shift was of approximately 0.5 km. Given that the methodology developed to map riverine tsetse densities in Burkina Faso was based on the land-use classification of a 1 km

[4] http://hydrosheds.cr.usgs.gov

FIGURE 6
(a) River ecotypes along the Mouhoun river (Burkina Faso), as derived from water surface analysis. (b) Levels of disturbance derived from swamp forest analysis. (c) ADT of *G. p. gambiensis*. (d) ADT of *G. tachinoides*



buffer around the river network, HydroSHEDS products could assist in applying this methodology at a broader, possibly regional, scale. Furthermore, HydroSHEDS provides drainage basins (watershed boundaries), which may prove useful in assisting the planning of tsetse genetic surveys. These surveys aim to identify isolated tsetse populations that can be targeted in area-wide control campaigns. It is believed that the potential of HydroSHEDS products to support tsetse and trypanosomiasis (T&T) studies and interventions should be further explored.

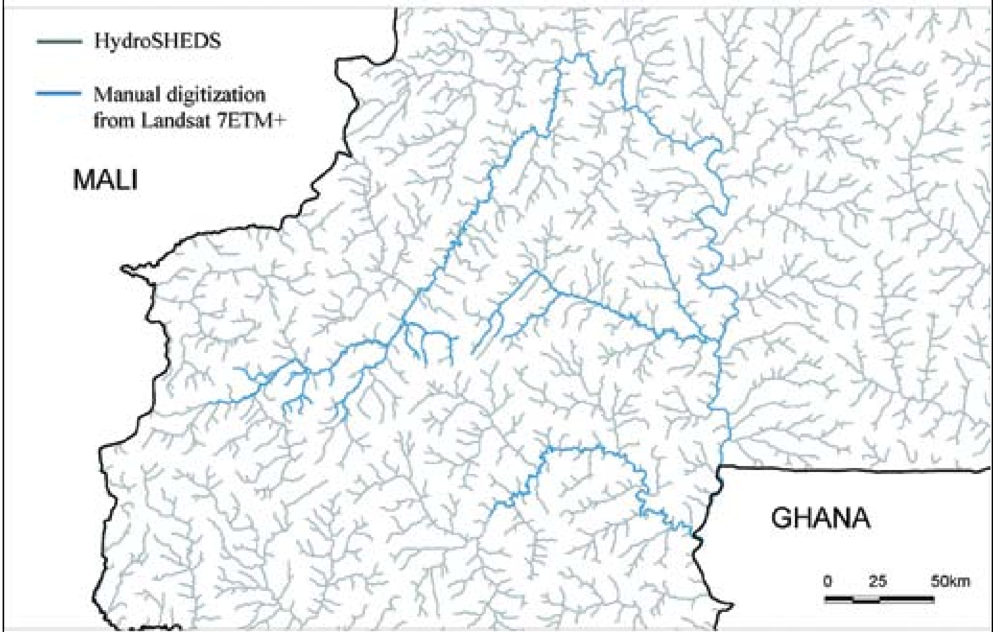
ANNUAL RAINFALL

We used the FAOClim[5] database (see page 24) to analyse the evolution of *G. p. gambiensis* and *G. tachinoides* distributional limits between entomological surveys

[5] http://www.fao.org/NR/climpag/pub/EN1102_en.asp

FIGURE 7

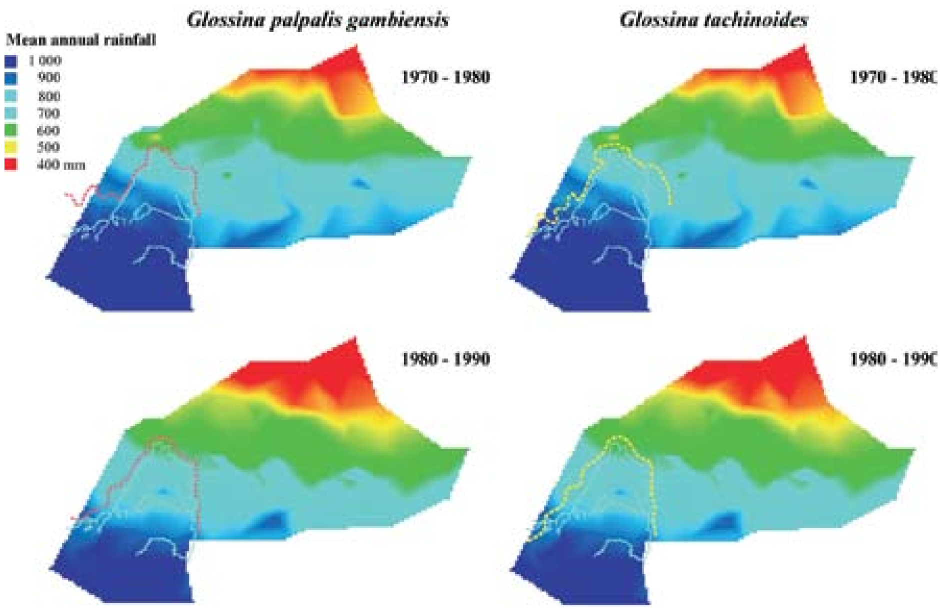
The Mouhoun river network, as digitized from Landsat 7 ETM+ images and the HydroSHEDS drainage network



carried out in 1979–1980 (Küpper, 1980) and in 1999–2007 (during various research programmes in which CIRDES was involved). One of the most important climatic factors used to map the density and distribution of tsetse flies is mean annual rainfall, which is positively correlated with indices of vegetation such as the normalized difference vegetation index (NDVI) (Rogers and Randolph, 1991). A map of the decadal mean annual rainfall for two periods (1970–1980 and 1980–1990) was interpolated from 151 meteorological stations in Burkina Faso and compared to the northern limit of tsetse distributions (Figure 8).

Although Figure 8 shows a clear tendency for the isohyets to move southward over time, the limits of the two tsetse species remain at the same latitude, at the top of the Mouhoun river loop. These limits thus seem to depend more on the persistence of the hydrological network than on the annual rainfall. At the same time, no tsetse are found beyond 800 mm of annual rainfall in the absence of permanent rivers or springs (i.e. east or west of the main course of the Mouhoun or springs). Thus, when integrated in multifactorial spatial models, the FAOClim database is likely to be useful in mapping riverine tsetse flies.

FIGURE 8
Decadal mean annual rainfall, as interpolated from 151 meteorological stations in the FAOClim database and northern borders of the distribution areas of *G. p. gambiensis* (left) and *G. tachinoides* (right)



CONCLUSION

A selection of the global GIS datasets presented in the first section of this publication (see page 1) was matched against local entomological and environmental datasets with a view to assessing their potential for supporting T&T interventions.

Datasets at a coarser resolution (the PAAT-IS tsetse distribution maps and GLC2000 land cover map of Africa) predictably showed their limits when utilized at a local scale, thus confirming the need for different approaches to support the planning and implementation of field interventions.

On the other hand, higher resolution global datasets available in the public domain (the NASA-LOIL and HydroSHEDS drainage network) demonstrated their potential for the estimation of disease risk. This potential merits further exploration in the context of T&T problem-solving.

REFERENCES

Bouyer, J., Guerrini, L., Cesar, J., De La Rocque, S. & Cuisance, D. 2005. A phyto-sociological analysis of the distribution of riverine tsetse flies in Burkina Faso. *Medical and Veterinary Entomology*, 19: 372–378.

Bouyer, J., Guerrini, L., Desquesnes, M., de La Rocque, S. & Cuisance, D. 2006. Mapping African animal trypanosomosis risk from the sky. *Veterinary Research,* 37(5): 633–645.

Cecchi, G., Mattioli, R.C., Slingenbergh, J. & de la Rocque, S. 2008. Land cover and tsetse fly distributions in sub-Saharan Africa. *Medical and Veterinary Entomology*, 22: 364–373. DOI: 10.1111/j.1365-2915.2008.00747.x.

de La Rocque, S., Michel, J.F., Bouyer, J., De Wispelaere, G. & Cuisance, D. 2005. Geographical information systems in parasitology: a review of potential applications using the example of animal trypanosomosis in West Africa. *Parasitologia,* 47: 97–104.

Guerrini, L. & Bouyer, J. 2007. Mapping African animal trypanosomosis risk: the landscape approach. *Veterinaria italiana*, 43(3): 643–654.

Guerrini, L. & Bouyer, J. In press. *A river-based model to predict riverine tsetse densities*. 29th Meeting of the International Scientific Council for Trypanosomiasis Research and Control (ISCTRC), Luanda, Angola.

Guerrini, L., Bord, J.P., Ducheyne, E. & Bouyer, J. 2008. Fragmentation analysis for prediction of suitable habitat for vectors: the example of riverine tsetse flies in Burkina Faso. *Journal of Medical Entomology,* 45(6): 1180–1186.

Hendrickx, G., Napala, A., Dao, B., Batawui, D., De Deken, R., Vermeilen, A. & Slingenbergh, J.H.W. 1999. A systematic approach to area-wide tsetse distribution and abundance maps. *Bulletin of Entomological Research*, 89: 231–244.

Itard, J., Cuisance, D. & Tacher, G. 2003. Trypanosomoses: historique – répartition géographique. In Tec & Doc, ed. *Principales maladies infectieuses et parasitaires du bétail. Europe et régions chaudes*, pp. 1607–1615. Paris, Editions Tec & Doc.

Küpper, W. 1980. Yahres und AbschluSbericht. Tsetse prospection OberVolta.1978–1980, p. 55, Bobo-Dioulasso.

Morel, P.C. 1983. *Guide pour la détermination des arbres et des arbustes dans les savanes Ouest-Africaines*. Maisons-Alfort, France, Institut d'Elevage et de Médecine Vétérinaire Tropicale.

Robinson, T., Rogers, D. & Brian, W. 1997. Mapping tsetse habitat suitability in the common fly belt of southern Africa using multivariate analysis of climate and remotely sensed vegetation data. *Medical and Veterinary Entomology*, 11: 235–245.

Rogers, D.J. & Randolph, S.E. 1991. Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature*, 351: 739–41.

Rogers, D.J. & Randolph, S.E. 1993. Distribution of tsetse and ticks in Africa: past, present and future. *Parasitology Today*, 9(7): 226–271.

Rogers, D.J., Hay, S.I. & Packer, M.J. 1996. Predicting the distribution of tsetse flies in West Africa using temporal Fourier-processed meteorological-satellite data. *Annals of Tropical Medicine and Parasitology*, 90: 225–241.