Patterning and predicting aquatic insect richness in four West-African coastal rivers using artificial neural networks

E.O. Edia[1], M. Gevrey[2], A. Ouattara[1], S. Brosse[2], G. Gourène[1], S. Lek[2]

ABSTRACT

Key-words:
*insect richness,
artificial
neural network,
prediction,
coastal rivers,
West Africa*

Despite their importance in stream management, the aquatic insect assemblages are still little known in West Africa. This is particularly true in South-Eastern Ivory Coast, where aquatic insect assemblages were hardly studied. We therefore aimed at characterising aquatic insect assemblages on four coastal rivers in South-Eastern Ivory Coast. Patterning aquatic insect assemblages was achieved using a Self-Organizing Map (SOM), an unsupervised Artificial Neural Networks (ANN) method. This method was applied to pattern the samples based on the richness of five major orders of aquatic insects (Diptera, Ephemeroptera, Coleoptera, Trichoptera and Odonata). This permitted to identify three clusters that were mainly related to the local environmental status of sampling sites. Then, we used the environmental characteristics of the sites to predict, using a multilayer perceptron neural network (MLP), trained by BackPropagation algorithm (BP), a supervised ANN, the richness of the five insect orders. The BP showed high predictability (0.90 for both Diptera and Trichoptera, 0.84 for both Coleoptera and Odonata, 0.69 for Ephemeroptera). The most contributing variables in predicting the five insect order richness were pH, conductivity, total dissolved solids, water temperature, percentage of rock and the canopy. This underlines the crucial influence of both instream characteristics and riparian context.

RÉSUMÉ

Assemblage et prédiction de la richesse taxonomique de l'entomofaune de quatre rivières de la côte ouest-africaine par des réseaux de neurones artificiels

Mots-clés :
*richesse
taxonomique
des insectes,
réseau de
neurones
artificiels,
rivières côtières,*

Malgré leur importance dans la gestion des hydrosystèmes, les insectes aquatiques restent peu connus en Afrique de l'Ouest et particulièrement dans le Sud-est de la Côte d'Ivoire. Cette étude vise donc à caractériser les assemblages des insectes aquatiques dans quatre rivières côtières du Sud-est ivoirien. La méthode non supervisée des réseaux de neurones artificiels, cartes auto-organisatrices (SOM), a été utilisée pour établir le patron de distribution des cinq ordres les plus diversifiés. Elle a permis d'identifier trois clusters qui sont principalement en relation avec les conditions environnementales des sites d'échantillonnage. L'algorithme de rétropropagation (BP), qui est un réseau de neurones supervisé, a servi

(1) Laboratoire d'Environnement et de Biologie Aquatique, U.F.R.-S.G.E., University of Abobo-Adjamé, 02 BP 801 Abidjan 02, Ivory Coast, oiedia@yahoo.fr
(2) Laboratoire Évolution Diversité Biologique, U.M.R. 5174, C.N.R.S. – University Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4, France

Afrique
de l'Ouest

à la prédiction de la richesse taxonomique des cinq ordres d'insectes à partir des caractéristiques environnementales des sites et a montré une forte prédictibilité (0,90 pour les Diptères et Trichoptères, 0,84 pour les Coléoptères et Odonates, 0,69 pour les Éphéméroptères). Les variables environnementales qui contribuent significativement à la prédiction de la richesse taxonomique des insectes aquatiques étaient le pH, la conductivité, le taux de solides dissous, la température de l'eau, le pourcentage de bloc et la canopée. Ces résultats traduisent l'effet combiné des caractéristiques des cours d'eau et de la zone riveraine sur cette communauté.

INTRODUCTION

In lotic ecosystems, the species composition of invertebrate community depends on the diversity and stability of stream habitats (Cummins, 1979) which provide ecological niches for the development of species (Malmqvist and Otto, 1987). Thus, understanding community patterns is important to manage target ecosystems. Therefore, macroinvertebrates are widely used as indicators of short- and long-term environmental changes in running waters (Hawkins *et al*., 2000). They provide both a facility for examining temporal changes and integrating the effects of prolonged exposure to intermittent discharges or variable concentrations of pollutants (Hellawell, 1986). Thus, it is promising to characterize the changes occurring in these macroinvertebrate communities to assess target ecosystems exposed to environmental disturbance.

Species richness is commonly used as an integrative descriptor of the community (Lenat, 1988). It is influenced by natural and/or anthropogenic disturbances (Rosenberg and Resh, 1993). Indeed, Resh and Jackson (1993) observed that species richness measures were sensitive to the impact of human activities on the stream ecosystems, particularly aquatic insects, which can be considered as good biological indicators of stream disturbances (Norris, 1995; Wright *et al*., 1995).

Despite their importance in stream ecosystems, aquatic insects are little studied in West Africa (Yaméogo *et al*., 2004). In Ivory Coast, only a few studies have been devoted to macroinvertebrate fauna (Dejoux *et al*., 1981; Sankaré, 1991; Diomandé, 2001; Diétoa, 2002; Edia *et al*., 2007). Among them, three were conducted in South-East of Ivory Coast. These studies only described the assemblage pattern of macroinvertebrate fauna of some rivers of this area (Diomandé, 2001; Diétoa, 2002; Edia *et al.*, 2007). In this work, we focussed on four small coastal rivers which play an important role for local human populations in South-East Ivory Coast, like most coastal rivers of Ivory Coast where water is used for domestic activities (drinking, cooking, bathing...), agriculture (irrigation, cattle drinking) and fisheries. It is therefore important to manage these streams in order to preserve water quality and the diversity of biological communities. Such management requires the knowledge of how environmental variables affect the richness and the taxonomic structure of these communities (Smogor and Angermeier, 1999).

Traditionally, conventional multivariate analyses have been applied to solve these problems (Legendre and Legendre, 1998). But using diverse species assemblages such as macroinvertebrates, we have to deal with ecological data that are bulky, nonlinear and complex, showing noise, redundancy, internal relations and outliers (Lek and Guégan, 2000). Artificial neural networks are known as well-adapted tools for dealing with such problems and to extract information out of complex, nonlinear data. These methods are more and more used for patterning assemblages (Worner and Gevrey, 2006; Brosse *et al*., 2007; de Thoisy *et al.*, 2008) and predicting occurrence and/or abundance of various taxa (Gevrey *et al*., 2005; Leprieur *et al*., 2006).

The aim of our study was first to pattern aquatic insect distribution so as to determine characteristic assemblages for the different sites and rivers. Then, it was to build a predictive model able to predict aquatic insect richness using environmental descriptors of the sites,

and finally determine the relative importance of environmental determinants on the richness and the composition of aquatic assemblages.

MATERIALS AND METHODS

> STUDY AREA

The study was undertaken in four coastal rivers located in the South-East of Ivory Coast: Soumié, Eholié, Ehania and Noé rivers (Figure 1). In each of these coastal rivers, two sampling sites were retained: one upstream and the other downstream (Figure 1). Table I summarizes environmental characteristics of these sites.

> AQUATIC INSECT AND ENVIRONMENTAL VARIABLE COLLECTION

Samples were collected during eight campaigns (between July 2003 and March 2005) at each sampling site between 08 and 10 a.m. Aquatic insects were sampled by means of drift net (mesh size: 250 μm) and hand net (mesh size: 250 μm). Drifting organisms were collected using a drift net suspended from a hand held rope.

For the hand net, samples were taken by submerging the net and sweeping it through the water column for a distance of 10 m. The net was also bumped against the bottom substrate to dislodge and collect organisms from the sediment. Two replicate samples were collected at each site and at each date. The samples were fixed in 10% formaldehyde. The three samples (1 by drift net and 2 by hand net) at each site and each campaign were pooled for analysis. In the laboratory, specimens were sorted and identified to the lowest possible taxonomic level by means of the keys in Dejoux *et al.* (1981), Day *et al.* (2002), de Moor *et al*. (2003a, 2003b), and Stals and de Moor (2007), and by consulting specialists.

At each campaign, each sampling site was characterized by measuring water temperature (TW), pH, conductivity (CND), total dissolved solids (TDS) and dissolved oxygen (DO) with portable sensors. Secchi disk transparency (Trans) measured using a 20-cm-diameter Secchi disk. Canopy and percentage of four substrata (rock, gravel, sand and clay/mud) were once visually assessed according to Gordon *et al.* (1994) methodology.

> MODELLING PROCEDURE

The modelling procedure was conducted in two steps. First, the Self-Organizing Map (SOM) (Kohonen, 1982) was used to cluster the samples according to species richness similarities. Secondly, a multilayer perceptron (MLP) using backpropagation algorithm (BP, Rumelhart *et al*., 1986) was used to predict the species richness (output variable) depending on environmental data (11 input variables).

The SOM neural network is composed of two layers of neurons: the input neurons linked to the samples and the output neurons which represent the map. This map consists of a certain number of neurons which are usually arranged into two dimensional grids in order to better visualization. In each of these output neurons there are virtual samples. The main concept of the SOM algorithm is based on a sequential computation of the Euclidean distance between samples and each virtual sample. At the end of the learning process, each sample is set in the neuron of the SOM for which the distance was the lowest. Neurons that are neighbours on the grid are expected to represent neighbour clusters of samples.

The detailed algorithm of the SOM can be found in Kohonen (2001) for theoretical considerations and in Chon *et al.* (1996), Lek and Guégan (1999) for ecological applications.

In this study, 64 samples were classified according to the number of macroinvertebrate orders per site and per campaign using the SOM. Five insect orders (Diptera, Ephemeroptera,
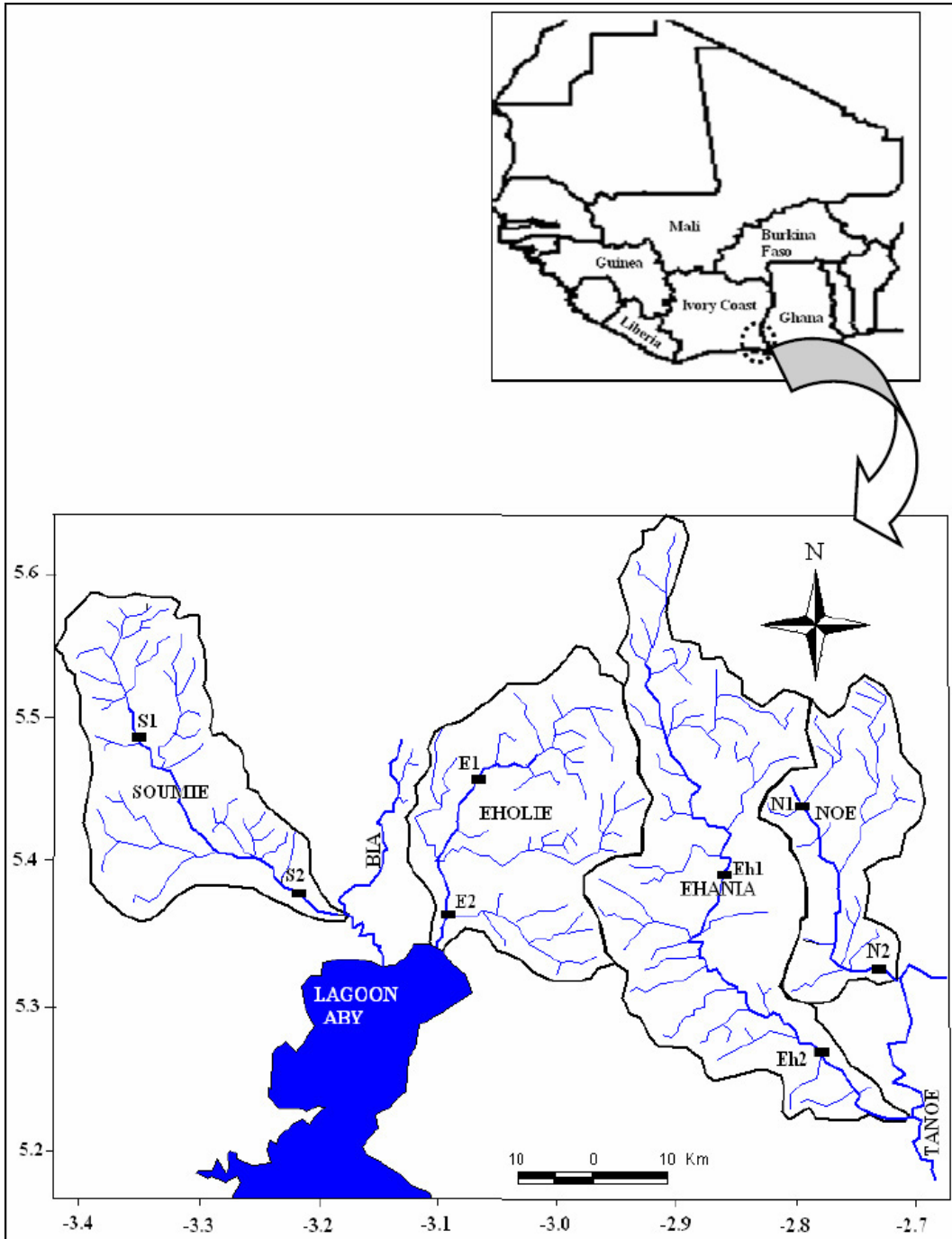
Figure 1
Location of the study area showing the four studied rivers. Dot marks indicate the sampling points on the four rivers. In station names, the letter indicates the river name (S: Soumié; E: Eholié; Eh: Ehania; N: Noé) and the number shows the station position on the river (1 = upstream and 2 = downstream).

Figure 1
Localisation de la zone d'étude montrant les quatre rivières étudiées. Les points noirs indiquent les sites d'échantillonnage sur les quatre rivières. Au niveau des noms des stations, la lettre indique le nom de la rivière (S : Soumié ; E : Eholié ; Eh : Ehania ; N : Noé) et le chiffre montre la position de la station sur la rivière (1 = amont et 2 = aval).

Table I
Characteristics and environmental variables (mean ± SE) of the eight study sites. Values without SE were estimated once.

Tableau I
Caractéristiques et variables environnementales (moyenne ± ES) des huit stations d'échantillonnage. Les valeurs sans ES ont été estimées une fois.

Parameters	Sampling sites							
	Soumié River		Eholié River		Ehania River		Noé River	
	Site S1	Site S2	Site E1	Site E2	Site Eh1	Site Eh2	Site N1	Site N2
Geographical positions	05° 29' N 03° 22' W	05° 24' N 03° 17' W	05° 28' N 03° 08' W	05° 23' N 03° 08' W	05° 24' N 02° 55' W	05° 17' N 02° 50' W	05° 28' N 02° 51' W	05° 18' N 02° 46' W
Water temperature (°C)	25.0 (0.38)	25.3 (0.41)	25.8 (0.52)	25.9 (0.37)	25.4 (0.29)	25.8 (0.38)	25.4 (0.23)	25.9 (0.39)
pH	7.1 (0.12)	6.8 (0.14)	7.0 (0.06)	7.0 (0.06)	7.1 (0.09)	6.9 (0.09)	7.0 (0.09)	6.7 (0.17)
Conductivity (µs·cm⁻¹)	57.4 (1.90)	42.6 (1.42)	55.2 (1.68)	60 (1.44)	64.1 (1.31)	54.9 (2.39)	64.2 (1.67)	54 (1.34)
Total dissolved solids (mg·L⁻¹)	27.3 (0.94)	20 (0.53)	25.8 (0.75)	26.8 (0.70)	29.6 (0.53)	25.8 (1.91)	30.1 (0.85)	25.1 (0.58)
Dissolved oxygen (mg·L⁻¹)	4.4 (0.34)	5.3 (0.49)	6 (0.62)	7.1 (0.76)	5.2 (0.54)	7.4 (0.50)	7.2 (0.49)	7.3 (0.56)
Secchi disk transparency (m)	0.61 (0.06)	0.60 (0.07)	0.52 (0.05)	0.53 (0.04)	0.74 (0.03)	0.50 (0.03)	0.58 (0.05)	0.44 (0.03)
Canopy (%)	35	55	70	85	40	45	5	10
% Rock in substrata	0	25	0	0	0	0	0	20
% Gravel in substrata	35	10	0	10	20	20	10	35
% Sand in substrata	45	40	30	50	20	45	40	25
% Clay/mud in substrata	20	25	70	40	60	35	50	20
Population density	Very low	Very low	Very low	High	Low	Very low	Low	High
Adjacent land use	Cultivated	Cultivated	Riparian forest	Habitations Cultivated	Habitations Cultivated	Riparian forest	Habitations Cultivated	Habitations

Very low: a few dispersed houses along the banks. Low: discontinuous habitat building along the banks. High: continuous habitat.

Coleoptera, Trichoptera and Odonata) were retained because they were major components of the aquatic insect fauna of the studied rivers (see Appendix).

A cluster analysis with the Ward algorithm can be used on the trained SOM to define the output neurons which are similar enough to be considered in the same cluster. To assess whether aquatic insect assemblages associated with each cluster were related to spatial factor (*i.e.* undisturbed and disturbed areas), we applied a proportion test based on χ^2 likelihood ratio statistics (*i.e.* G-test with Yates' correction, Zar, 1999). This test was regarded as significant at the level $p < 0.05$.

The BP structure was made up of a three-layered feed-forward network with information flowing from the input to the output layer. The input nodes comprised the environmental variables recorded per site and per campaign (64 samples) and the output node corresponded to the genus richness of each order. The information flow was represented by changing weights. A typical sigmoid function was used as the transfer function. Training was carried out by changing the weights according to the prediction errors until the generalisation error reached the global minimum (*i.e.* early-stopping). As the number of observations was not enough for splitting the data into training and testing sets, the BP was instead trained with the leave-one-out method (Kohavi, 1995). It consists of considering each observation (*i.e.* a sample) as a unique piece of information, repeating the estimation $n - 1$ times. The leave-one-out method is testing technique frequently employed in ecology (Lek *et al*., 2000).

The BP algorithm was written by the authors using Matlab® platform. The detailed algorithm of the BP can be found in Rumelhart *et al.* (1986) for theoretical considerations and in Lek and Guégan (1999) for ecological applications.

An important issue in model evaluation is determining the relative contribution (*i.e.* explanatory importance) of each predictive variable. To do this, the partial derivatives (PaD) algorithm was used (Gevrey *et al*., 2003). Five hundred models were set up for each of the five studied aquatic insect orders. The average and the standard error of the contribution were calculated for each environmental variable. In this study, an environmental descriptor was regarded as significant to the model when its contribution was at least 10% (Brosse and Lek, 2002).

RESULTS

> CLASSIFICATION OF SAMPLES USING SOM

The non-linear projection of the dataset in two dimensional space allowed us the classification of our samples according to Diptera, Ephemeroptera, Coleoptera, Trichoptera and Odonata richness (Figures 2A and 2B).

The units of the SOM map were classified into three different groups (I, II and III) based on the cluster analysis with Ward algorithm (Figure 2B). The grey scale displays different clusters on the SOM map (Figure 2A). Clusters I and II mainly consisted of samples from sites which are relatively exempt from disturbance (S1, S2, E1, Eh2) and cluster III of sites (E2, Eh1, N1, N2) which are the most disturbed by anthropogenic activities, such as agricultural and domestic activities, as they are located close to populated areas. The test of proportion confirmed this result. Indeed, samples from clusters I and II were significantly related to relatively undisturbed areas, whereas those from cluster III were significantly related to relatively disturbed areas.

Figure 3 displays component of each order richness (Diptera, Ephemeroptera, Coleoptera, Trichoptera and Odonata) in each neuron on the trained map in grey scale. Dark areas represent high richness, while light ones reveal low richness. Overall, bottom areas of the SOM map showed the highest aquatic insect richness, whereas top areas had low richness. Samples in cluster I had high Diptera and Odonata richness with moderate richness for other insects. Samples in cluster II had high numbers of Ephemeroptera, Coleoptera and Trichoptera genera, along with moderate richness for Diptera and low richness for Odonata. Cluster III had moderate richness of Diptera and Odonata with low richness for other taxa.
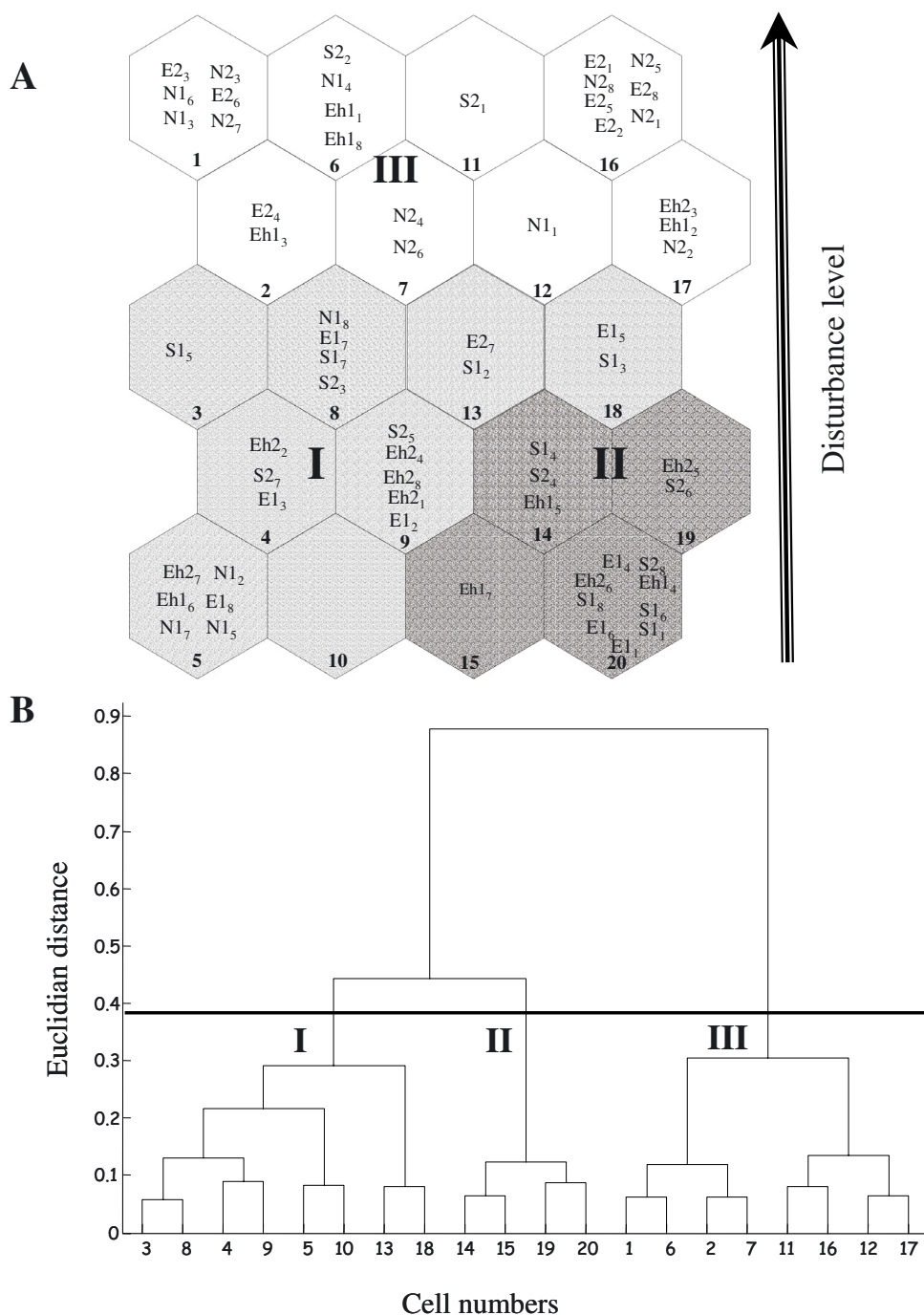
Figure 2

A) Samples classification according to Diptera, Ephemeroptera, Coleoptera, Trichoptera and Odonata genus richness, using the SOM. The latin numbers (I–III) represent different clusters. The arabic numbers (1–20) represent the SOM units. The arrow indicates the level of disturbance. B) Hierarchical classification of SOM units using Ward's algorithm. The numbers correspond to the cell numbers of the map in Figure 2A.

Figure 2

A) Classification des échantillons sur la base de la richesse générique des Diptères, Ephéméroptères, Coléoptères, Trichoptères et Odonates à partir de la SOM. Les chiffres romains (I–III) représentent les différents clusters. Les chiffres arabes (1–20) représentent les unités de la SOM. La flèche indique le degré de perturbation des stations. B) Classification hiérarchique des unités de la SOM à partir de la méthode Ward. Les chiffres correspondent aux numéros des cellules de la carte sur la Figure 2A.
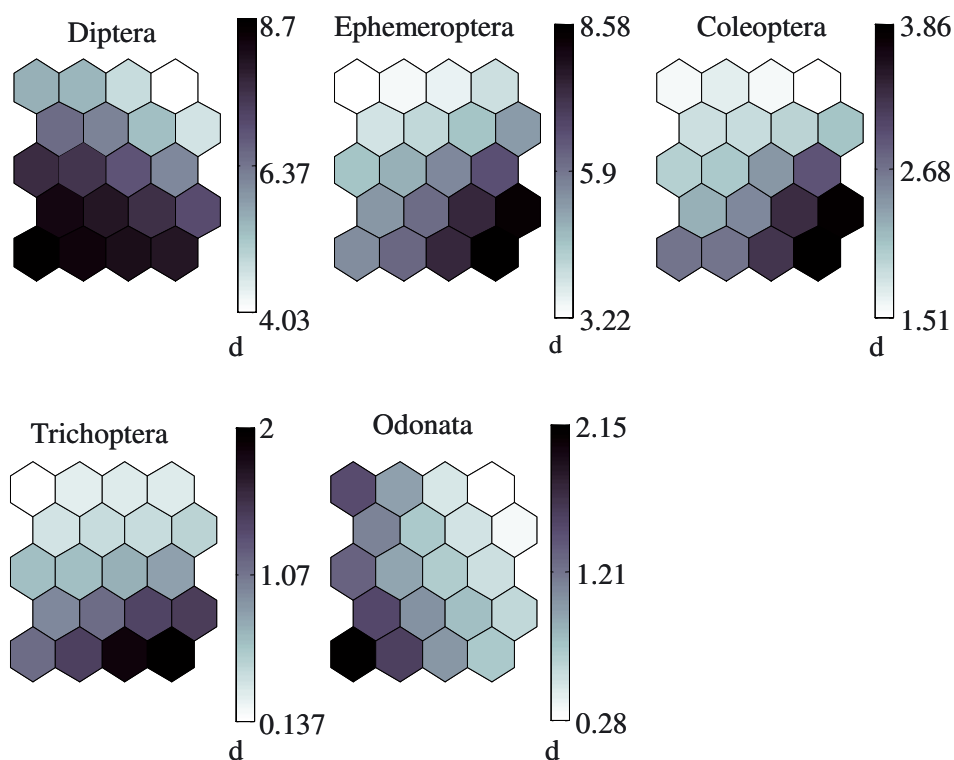
Figure 3
Visualization of each order richness calculated in the trained SOM in grey scale. Dark represents high value richness, while light is low value richness.

Figure 3
Visualisation de la richesse taxonomique de chaque ordre calculée par la SOM. Les parties sombres représentent une richesse élevée ; les claires indiquent les faibles richesses.

> PREDICTION OF AQUATIC INSECT RICHNESS

The results from BP training showed accuracy in predicting the overall aquatic insect richness on the basis of the environmental variables. Indeed, on the set of the 64 samples, correlation coefficient between observed and estimated values was 0.90 for both Diptera and Trichoptera richness. There was also high correlation between observed and estimated values (0.84) both for Coleoptera and Odanata. For Ephemeroptera, the correlation coefficient between observed and estimated values was 0.69.

In terms of ecological information extracted from the BP model, the result of the PaD algorithm stress the importance of the environmental variables in the prediction of the five aquatic insect orders. The importance of contribution in percentage for each descriptor after applying the partial derivatives method is given in Figure 4. Five parameters governed the Diptera richness: pH, conductivity, dissolved oxygen, percentage of rock and canopy. With exception of pH, these variables were turned out to be significant in predicting Ephemeroptera richness. However, the conductivity was the most contributing parameter in the prediction of this order. Moreover, in decreasing order, dissolved oxygen, transparency, total dissolved solids, and percentage of rock were meaningful in the Coleoptera prediction. In the same order, the main variables which determine the Trichoptera richness were water temperature, conductivity and transparency. Concerning the Odonata richness, pH was the most contributing variable. It was followed by water temperature, canopy, conductivity and total dissolved solids.
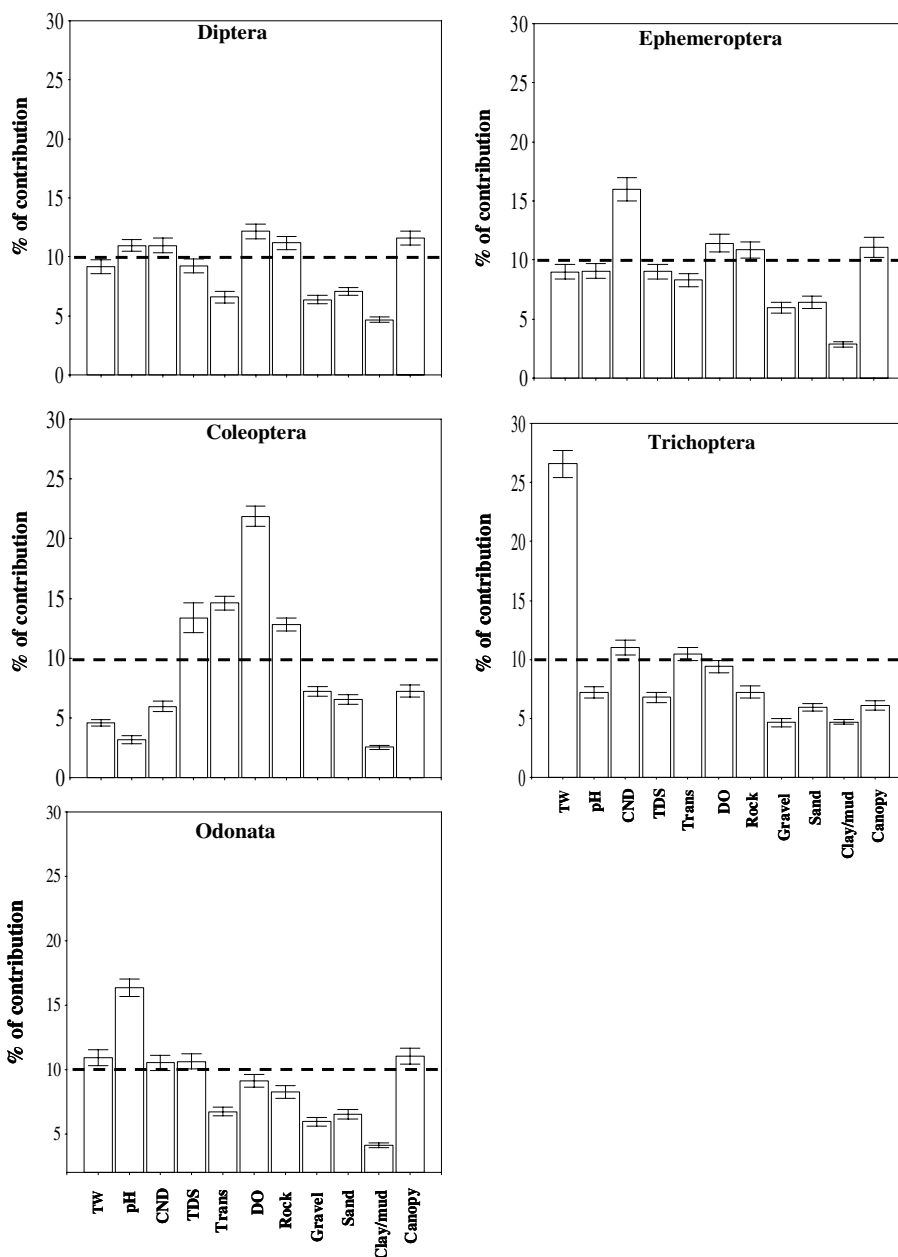
Figure 4

Percentage contributions of each of the eleven independent variables to the prediction of each aquatic insect order, obtained by partial derivatives algorithm. The dotted line represents the significance level (i.e. 10%). The environmental variables are: water temperature (TW), pH, conductivity (CND), total dissolved solids (TDS), transparency (Trans), dissolved oxygen (DO), canopy and percentages of four substrata: rock, gravel, sand and clay/mud. Bars indicate the mean value of the results of the 500 models for each order; horizontal lines represent standard errors of the mean.

Figure 4

Pourcentage de contribution des variables environnementales dans la prédiction de chaque ordre d'insectes aquatiques obtenu par l'algorithme des dérivées partielles. La ligne en pointillés indique le niveau de significativité des contributions (*i.e.* 10 %). Les variables environnementales sont la température de l'eau (TW), le pH, la conductivité (CND), le taux de solides dissous (TDS), l'oxygène dissous (DO), la canopée et les pourcentages des quatre substrats : roche, gravier, sable, argile/limon. La barre indique la valeur moyenne des résultats des 500 modèles pour chaque ordre ; les lignes horizontales représentent les erreurs standard de la moyenne.

DISCUSSION

Samples clustering by the SOM can mainly be afforded to the impact of anthropogenic activities. For example, the samples gathered in cluster III are mainly from the sites which are the most disturbed by anthropogenic activities (agricultural and domestic activities), as they are located close to populated areas. On the other hand, samples gathered in clusters I and II are mainly from the sites which are relatively exempt from disturbance. These anthropogenic disturbances affect some environmental variables such as conductivity and total dissolved solids in these areas and potentially reduce the aquatic insect diversity (Azrina *et al.*, 2006).

The supervised artificial neural network (MLP) showed a high accuracy in predicting the overall orders richness on the basis of the environmental variables. This accuracy may be explained by ability of the MLP to deal with nonlinear and complex data such as ecological data (Lek *et al.*, 1996; Gevrey *et al.*, 2005).

Moreover, the MLP model, with help of sensitivity analysis, revealed the relative importance of each environmental variable in structuring aquatic insect richness. The most contributing variables in predicting aquatic insect richness can be gathered into chemical variables (pH, conductivity, total dissolved solids and dissolved oxygen) and physical parameters (water temperature, percentage of rock and canopy). Our results were ecologically congruent. Indeed, the pH, the conductivity and the total of dissolved solids which express water mineralization level are linked to nutrient accumulation (Galdean and Staicu, 1997). The nutrient accumulation which is related to the landscape disturbance (Freeman *et al*., 2007), is known to control the dissolved oxygen ratio through the algal growth (Turner and Rabalais, 1994; Justic *et al*., 1995). The high nutrient accumulation in aquatic systems is accompanied by the increase of acidity, conductivity and total dissolved solids and by the decrease of dissolved oxygen. This accumulation affects energy flow of aquatic systems and causes the decline of biodiversity (Wiederholm, 1984; Schindler, 1990). Concerning the physical variables, the contribution of water temperature in predicting aquatic insect richness was, in ecological view, relevant. Indeed, water temperature is an important indicator of watershed vitality because of its controlling influence on the metabolism, development and activity of aquatic organisms (Naiman *et al*., 1992). The importance of the percentage of rock and the canopy in the model was also reliable. The presence of rock in streambed offers potential refugia to macroinvertebrates (Gayraud, 2001). This fact limits the organism drift able to modify the macroinvertebrate assemblage (Townsend, 1989). Moreover, according to Matthaei *et al*. (1999), streambed instability is known to influence macroinvertebrate diversity. It is recognized that the riparian vegetation plays a main rule in food availability in tropical streams (Benstead *et al*., 2003). Moreover, plant communities in riparian areas help to determine what, how much, and when materials from upland areas enter the stream ecosystem. For example, a wide, mature riparian forest will capture many soils and sediments, nutrients, and woody debris, adding richness and complexity to soil and plant communities near the water and protecting water from excessive nutrient or soil inputs (Lowrance *et al*., 1986). So, riparian areas and consequently the structure, functions and processes occurring within and around the stream, are fundamentally altered when significant upland and riparian vegetation is removed.

The Rapid Bioassessment Protocols are being applied in many countries with success, optimizing time and resources in sample methodologies. Unfortunately, in tropical developing countries like Ivory Coast, there is a lack of basic knowledge on taxonomics and distribution patterns for macroinvertebrates. Thus, more studies are needed to understand the tropical benthic community structure better. Afterwards, we will probably be able to adapt and develop our own biotic indices and many other important metrics for water quality evaluation.

REFERENCES

Azrina M.Z., Yap C.K., Rahim Ismail A., Ismail A. and Tan S.G., 2006. Anthropogenic impacts on the distribution and biodiversity of benthic macroinvertebrates and water quality of the Langat River, Peninsular Malaysia. *Ecotox. Environ. Safe.*, *64*, 337–347.

Benstead J.P., De Rham P.H., Gattolliat J.-L., Gibon F.-M., Loiselle P.V., Sartori M., Sparks J.S. and Stiassny M.L.J., 2003. Conserving Madagascar's freshwater biodiversity. *Bioscience*, *53*, 1101–1111.

Brosse S. and Lek S., 2002. Relationship between environmental characteristics and the density of age-0 Eurasian Perch *Perca fluviatilis* in littoral zone of lake: a nonlinear approach. *T. Am. Fish. Soc.*, *131*, 1033–1043.

Brosse S., Grossman G.D. and Lek S., 2007. Fish assemblage patterns in the littoral zone of a European reservoir. *Freshwater Biol.*, *52*, 448–458.

Chon T.S., Park Y.S., Moon K.H., Cha E. and Pa Y., 1996. Patternizing communities by using an artificial neural network. *Ecol. Model.*, *90*, 69–78.

Cummins K.W., 1979. The natural stream ecosystem. *In*: Ward J.V. and Standford J.A. (eds.), The Ecology of Regulated Streams, Plenum Press, New York, 7–24.

de Moor I.J., Day J.A. and de Moor F.C., 2003a. Guides to the Freshwater Invertebrates of Southern Africa, Volume 7, Insecta I: Ephemeroptera, Odonata and Plecoptera, WRC report N° TT 207/03, South Africa, 288 p.

de Moor I.J., Day J.A. and de Moor F.C., 2003b. Guides to the Freshwater Invertebrates of Southern Africa. Volume 8, Insecta II: Hemiptera, Megaloptera, Neuroptera, Trichoptera and Lepidoptera, WRC report N° TT 214/03, South Africa, 209 p.

de Thoisy B., Brosse S. and Dubois M., 2008. Assessment of large-vertebrate species richness and abundance in neotropical forest using line-transect censuses: what is the minimal effort required? *Biodivers. Conserv.*, *17*, 2627–2644.

Day J.A., Harrison A.D. and de Moor I.J., 2002. Guides to the Freshwater Invertebrates of Southern Africa, Volume 9: Diptera, WRC report N° TT 201/02, South Africa, 200 p.

Dejoux C., Elouard J.M., Forge P. and Maslin J.L., 1981. Catalogue Iconographique des Insectes Aquatiques de Côte d'Ivoire, Report ORSTOM, Bouaké, Ivory Coast, 172 p.

Diétoa Y.M., 2002. Entomofaune et stratégies alimentaires des poissons du genre *Brycinus* (Characidae) en milieux fluviatiles et lacustres (Bassins Bia et Agnéby ; Côte d'Ivoire). Ph.D. thesis, University of Abobo-Adjamé, Abidjan, Ivory Coast, 261 p.

Diomandé D., 2001. Macrofaune benthique et stratégies alimentaires de *Synodontis bastiani* et *S. schall* en milieu fluvio-lacustre (Bassins Bia et Agnébi, Côte d'Ivoire). Ph.D. thesis, University of Abobo-Adjamé, Abidjan, Ivory Coast, 260 p.

Edia O.E., Brosse S., Ouattara A., Gourène G., Winterton P. and Lek-Ang S., 2007. Aquatic insect assemblage patterns in four West-African coastal rivers. *J. Biol. Sci.*, *7*, 1130–1138.

Freeman M.C., Pringle C.M. and Jackson C.R., 2007. Hydrologic connectivity and contribution of stream headwaters in ecological integrity at regional scales. *J. Am. Water Resour. As.*, *43*, 5–14.

Galdean N. and Staicu G., 1997. The carrying capacity assessment of the lotic system Crisul Repede (Tisa Area Catchment, Romania), based on faunistical analysis. *Trav. Mus. Natl. Hist. Nat. 'Grigore Antipa'*, *37*, 237–254.

Gayraud S., 2001. Les sédiments du lit des cours d'eau : quantification de leur influence sur les peuplements de macroinvertébrés par une approche multi-sites. Ph.D. thesis, University Claude Bernard, Lyon, France, 188 p.

Gevrey M., Dimopoulos I. and Lek S., 2003. Review and comparison of methods to study the contribution of variables in Artificial Neural Network models. *Ecol. Model.*, *160*, 249–264.

Gevrey M., Park Y.S., Verdonschot P.F.M. and Lek S., 2005. Predicting Dutch macroinvertebrate species richness and functional feeding groups using five modelling techniques. *In*: Lek S., Scardi M., Verdonschot P.F.M., Descy J.P. and Park Y.S. (eds.), Modelling Community Structure in Freshwater Ecosystems, Springer, Berlin, 158–166.

Gordon N.D., McMahon T.A. and Finlayson B.L., 1994. Stream Hydrology, an Introduction for Ecologists, John Wiley and Sons, New York, 526 p.

Hawkins C.P., Norris R.H., Hogue J.N. and Feminela J.W., 2000. Development and evaluation of predictive models for measuring the biological integrity. *Ecol. Appl.*, *10*, 1456–1477.

Hellawell J.M., 1986. Biological Indicators of Freshwater Pollution and Environmental Management, Elsevier, London, 546 p.

Justic D., Rabelais N.N., Turner R.E. and Dortch Q., 1995. Changes in nutrient structure of river-dominated coastal waters: stoichiometric nutrient balance and its consequences. *Est. Coast. Shelf Sci.*, *40*, 339–356.

Kohavi R., 1995. A study of cross-validation and bootstrap for estimation and model selection. *In*: 14th International Joint Conference on Artificial Intelligence, Montréal, Canada, 1137–1143.

Kohonen T., 1982. Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, *43*, 59–69.

Kohonen T., 2001. Self-Organizing Maps, 3rd edition, Springer, Berlin, 501 p.

Legendre P. and Legendre L., 1998. Numerical ecology, Elsevier, Amsterdam, 853 p.

Lek S. and Guégan J.F., 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.*, *120*, 65–73.

Lek S. and Guégan J.F., 2000. Artificial Neuronal Networks: Application to Ecology and Evolution, Springer-Verlag, Berlin, 262 p.

Lek S., Delacoste M., Baran P., Dimopoulos J., Lauga J. and Aulagnier S., 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecol. Model.*, *90*, 39–52.

Lek S., Giraudel J.L. and Guégan J.F., 2000. Neuronal networks: algorithms and architectures for ecologists and evolutionary ecologists. *In*: Lek S. and Guégan J.P. (eds.), Artificial Neuronal Networks: Application to Ecology and Evolution, Springer-Verlag, Berlin, 3–27.

Lenat D.R., 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. *J. North Am. Benthol. Soc.*, *7*, 222–233.

Leprieur F., Hickey M.A., Arbuckle C.J., Closs G.P., Brosse S. and Townsend C.R., 2006. Hydrological disturbance benefits a native fish at the expense of an exotic fish. *J. Appl. Ecol.*, *43*, 930–939.

Lowrance R., Sharpe J.K. and Sheridan J.M., 1986. Long-term sediment deposition in the riparian zone of a coastal plain watershed. *J. Soil Water Conserv.*, *41*, 266–271.

Malmqvist B. and Otto C., 1987. The influence of substrate stability on the composition of stream benthos: an experimental study. *Oikos*, *48*, 33–38.

Matthaei C.D., Peacock K.A. and Townsend C.R., 1999. Patchy surface stone movement during disturbance in a New Zealand stream and its potential significance for fauna. *Limnol. Oceanogr.*, *44*, 1091–1102.

Naiman R.J., Beechie T.J., Benda L.E., Berg D.R., Bisson P.A., MacDonald L.H., Connor M.D.O., Olson P.L. and Steel E.A., 1992. Fundamental elements of ecologically healthy watersheds in the Pacific Northwest coastal ecoregion. *In*: Naiman R.J. (ed.), Watershed Management: Balancing Sustainability and Environmental Change, Springer-Verlag, New York, 127–188.

Norris R.H., 1995. Biological monitoring: the dilemma of data analysis. *J. North Am. Benthol. Soc.*, *14*, 440–450.

Resh V.H. and Jackson J.K., 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. *In*: Rosenberg D.M. and Resh V.H. (eds.), Freshwater Biomonitoring and Benthic Macroinvertebrate, Chapman and Hall, London, 195–223.

Rosenberg D.M. and Resh V.H., 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates, Chapman & Hall, New York, London, 488 p.

Rumelhart D.E., Hinton G.E. and Williams R.J., 1986. Learning internal representations by error propagation. *In*: Rumelhart D.E. and McCelland J.L. (eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Foundations, MIT Press, Cambridge, 318–362.

Sankaré Y., 1991. Étude comparative de la macrofaune associée aux racines de *Pistia stratiotes* L. (Araceae) du lac de barrage d'Ayamé II et du fleuve Comoé (Côte d'Ivoire). *J. Ivoir. Océanol. Limnol.*, *1*, 131–138.

Schindler D.W., 1990. Experimental perturbations of whole lakes as tests of hypotheses concerning ecosystem structure and function. *Oikos*, *57*, 25–41.

Smogor R.A. and Angermeier P.L., 1999. Effects of drainage basin and anthropogenic disturbance on relations between stream size and IBI metrics in Virginia. *In*: Simon T.P. (ed.), Assessment Approaches for Estimating Biological Integrity Using Fish Assemblages, Lewis Press, Boca Raton, 249–272.

Stals R. and de Moor I.J., 2007. Guides to the Freshwater Invertebrates of Southern Africa, Volume 10: Coleoptera, WRC report N° TT 320/07, South Africa, 263 p.

Townsend C.R., 1989. The patch dynamic concept of stream community ecology. *J. North Am. Benthol. Soc.*, *8*, 36–50.

Turner R.E. and Rabalais N.N., 1994. Coastal eutrophication near the Mississippi River delta. *Nature*, *368*, 619–621.

Wiederholm T., 1984. Responses of aquatic insects to environmental pollution. *In*: Resh V.H. and Rosenberg D.M. (eds.), The Ecology of Aquatic Insects, Praeger Pubs., New York, 508–557.

Worner S.P. and Gevrey M., 2006. Modelling global insect pest species assemblages to determine risk of invasion. *J. Appl. Ecol.*, *43*, 858–867.

Wright J.F., Furse M.T. and Armitage P.D., 1995. R1VPACS: A technique for evaluating the biological quality of rivers in the UK. *Eur. Water Pollut. Contr.*, *3*, 15–25.

Yaméogo L., Resh V.H. and Molyneux D.H., 2004. Control of rivers blindness in West Africa: Case history of biodiversity in a disease control program. *EcoHealth*, *1*, 172–183.

Zar J.H., 1999. Biostatistical Analysis, 4th edition, NJ Prentice Hall, Englewoods Cliffs, New Jersey, 662 p.

Appendix. List of the aquatic insect taxa found at the eight sampling sites. * Indicates the presence of taxa.

Appendice. Liste des insectes aquatiques collectés aux huit stations. * Indique la présence des taxons.

			Soumié		Eholié		Ehania		Noé	
Orders	Families	Taxa	S1	S2	E1	E2	Eh1	Eh2	N1	N2
Collembola	Arthropleona			*				*	*	*
Ephemeroptera	Leptophlebiidae	*Adenophlebiodes* sp.	*	*	*	*	*	*		
		Choroterpes sp.	*	*	*	*	*	*		*
		Euthraulus sp.	*	*			*	*		*
		Hyalophlebia sp.				*				
		Thraulus sp.	*	*		*	*	*		*
	Tricorythidae	*Dicercomyzon* sp.	*	*	*	*	*	*	*	*
		Tricorythus sp.	*	*	*	*	*	*		*
	Machadorythidae	*Machadorythus maculatus*			*	*				
	Ephemerythidae	*Ephemerythus* sp.	*	*			*	*		
	Polymitarcyidae	*Ephoron* sp.						*		*
	Caenidae	*Caenis* sp.	*	*	*	*	*	*	*	*
	Baetidae	*Afrobaetodes* sp.			*		*			*
		Bugilliesia sp.		*						
		Cloeodes dentatus				*				
		Cloeon sp.							*	*
		Cheleocloeon yolandae	*		*			*	*	
		Dabulamanzia babaora	*	*						
		Labiobaetis gambiae	*	*	*	*	*	*	*	*
		Procloeon sylvicola	*	*	*	*	*	*	*	*
		Susua sp.	*	*					*	
	Oligoneuriidae	*Elassoneuria* sp.			*			*		
	Heptageniidae	*Afronurus* sp.	*	*	*	*	*	*	*	*
		Composoneuria njalensis	*	*	*	*	*	*	*	*
		Notonurus sp.	*	*		*	*	*		
Plecoptera	Perlidae	*Neoperla* sp.					*			
Odonata	Calopterygidae	*Phaon iridipennis*					*			
	Coenagrionidae	*Coenagrion* sp.	*	*	*		*	*	*	
		Pseudogrion sp.				*				
	Gomphidae	*Lestinogomphus angustus*		*		*		*	*	
		Microgomphus sp.		*	*					
		Paragomphus sp.	*				*	*	*	
		Phyllogomphus aethiops		*	*	*	*	*	*	
	Cordulegasteridae	*Cordulegaster* sp.	*							
	Libellulidae	*Libellula* sp.			*	*				
		Olpogastra sp.					*	*		
		Zygonyx sp.		*						
		Palpopleura sp.					*			
	Macromiidae	*Macromia* sp.	*	*	*		*	*	*	*
		Phyllomacromia sp.	*	*	*	*	*		*	
	Chlorocyphidae	*Chlorocypha* sp.					*			
	Pleidae	*Plea* sp.		*		*	*		*	
	Notonectidae	*Anisops* sp.		*	*			*	*	
	Corixidae	*Micronecta scutellaris*	*		*					*
	Hydrometridae	*Hydrometra* sp.					*		*	
	Veliidae	*Microvelia* sp.		*			*	*	*	
Heteroptera	Veliidae	*Rhagovelia reitteri*	*	*	*		*	*		
	Gerridae	*Eurymetra* sp.			*	*	*	*	*	
		Gerris sp.	*							
	Belostomatidae	*Diplonychus* sp.			*					
		Limnogeton fieberi					*			
Lepidoptera	Crambidae					*	*	*	*	*
Hymenoptera			*	*			*	*		
Coleoptera	Gyrinidae	*Orectogyrus* sp.	*	*				*		
	Dytiscidae	*Copelatus* sp.							*	
		Dytiscus sp.	*		*		*		*	
		Laccophilus sp.	*	*	*		*	*	*	

Appendix. (Continued.)
Appendice. (Suite.)

Orders	Families	Taxa	Soumié S1	Soumié S2	Eholié E1	Eholié E2	Ehania Eh1	Ehania Eh2	Noé N1	Noé N2
	Hydrophilidae	*Enochrus* sp.						*		*
		Hydrobius sp.		*						*
	Elmidae	*Potamophilus* sp.					*			
		Potamodytes sp.		*			*		*	
		Elmis sp.	*	*	*	*	*	*	*	*
		Esolus sp.		*	*	*	*	*	*	*
		Limnius sp.	*	*	*	*	*	*	*	*
		Normandia sp.	*	*	*				*	
		Riolus sp.	*	*	*	*	*	*	*	*
		Dupophilus sp.		*	*	*	*	*	*	
		Oulimnius sp.	*		*	*	*			
		Macronychus sp.	*				*	*		
	Helodidae		*							
	Hydroscaphidae	*Hydroscapha* sp.		*						
Trichoptera	Hydropsychidae	*Cheumatopsyche* sp.	*	*				*	*	
		Polymorphanisus sp.								*
	Polycentropodidae	*Neureclipsis* sp.				*				
	Ecnomidae	*Ecnomus* sp.					*	*		
	Hydroptilidae	*Afritrichia* sp.	*				*	*	*	
		Hydroptila sp.			*		*		*	
		Orthotrichia sp.		*	*	*			*	
	Leptoceridae	*Ceraclea* sp.		*	*	*	*	*		
		Leptocerus sp.	*				*	*		
		Oecetis sp.	*	*	*		*	*	*	
		Triaenodes sp.		*		*				*
		Parasetodes sp.	*	*	*					
Diptera	Psychodidae									*
	Ptychopteridae	*Ptychopteria* sp.								*
	Chaoboridae	*Chaoborus* sp.								*
	Culicidae	*Aedes* sp.					*			
		Anopheles sp.	*	*	*		*	*	*	
		Culex sp.				*				
		Culicinae	*							
	Simuliidae	*Simulium damnosum*	*	*		*	*	*	*	*
	Ceratopogonidae	*Ceratopogon* sp.	*	*	*	*	*	*	*	
		Dasyheleinae							*	
		Forcipomyinae						*		
	Chironomidae	*Ablabesmyia* sp.	*	*	*	*	*	*	*	*
		Chironomus sp.	*	*	*	*	*	*	*	*
		Clinotanypus claripennis		*	*	*	*	*	*	*
Diptera	Chironomidae	*Cricotopus* sp.	*	*	*	*	*	*	*	*
		Cryptochironomus sp.	*	*	*	*	*	*	*	*
		Lauterborniella sp.						*		
		Nanocladius sp.	*	*	*	*	*	*	*	*
		Nilodorum sp.			*	*	*	*	*	*
		Orthocladiinae	*		*				*	
		Polypedilum sp.	*	*	*	*	*	*	*	*
		Procladius sp.				*				
		Stenochironomus sp.	*		*	*	*			
		Stictochironomus sp.	*	*	*	*	*	*	*	*
		Tanypus sp.				*	*		*	*
		Tanytarsus sp.	*	*	*	*	*	*	*	*
	Stratiomyidae								*	
	Empididae	Hemerodromiinae							*	*
	Athericidae	*Atherix* sp.	*		*	*	*			*
	Anthomyidae			*						
	Tabanidae	*Tabanus* sp.				*	*	*		*
	Tipulidae							*	*	