
Design and application of a stratified sampling strategy to study the regional distribution of cyanobacteria (Ile-de-France, France)

Arnaud Catherine^{a,*}, Marc Troussellier^b and Cécile Bernard^a

^a Unité Ecosystèmes et interactions toxiques, USM 0505/EA 4105 MNHN, Muséum National d'Histoire Naturelle, Case 39, 12 rue Buffon, F-75231 Paris Cedex 5, France

^b Laboratoire Ecosystèmes lagunaires, UMR 5119 Université Montpellier 2 – CNRS – Ifremer, Université Montpellier 2, Case 93, Place E. Bataillon, F-34095 Montpellier Cedex 5, France

*: Corresponding author : Arnaud Catherine, email address : arnocat@mnhn.fr

Abstract:

This study describes the design and application of a stratified sampling strategy of waterbodies to assess and analyze the distribution of cyanobacteria at a regional scale (Ile-de-France, IDF). Ten groups of hydrographical zones were defined within the IDF on the basis of their anthropogenic and geomorphologic characteristics. Sampling effort ($n = 50$) was then randomly allocated according to the number of waterbodies in each group. This sampling strategy was tested in August 2006, using a field probe to estimate total phytoplankton as well as cyanobacteria biomasses. The sampled waterbodies exhibited a wide range of phytoplankton ($<1-375 \mu\text{g equiv. Chl}a \text{ L}^{-1}$) and cyanobacteria biomasses ($<1-278 \mu\text{g equiv. Chl}a \text{ L}^{-1}$). 72% of the waterbodies in the IDF were classified as eutrophic (42% hypereutrophic), and 24% of the sites studied were dominated by cyanobacteria. Waterbodies connected to hydrographical networks ($n = 26$) showed significantly higher total ($p < 0.0001$; 3.5 times greater) and cyanobacterial ($p < 0.001$, 3.2 times greater) biomasses than the isolated ones ($n = 24$). No significant overall relationship was found through contingency analysis between waterbody trophic status and global land use categories (urban, periurban, and rural) within their hydrographical zones. However, concerning the waterbodies linked to hydrographical networks, the percentage of land covered by forest appeared as a good indicator of phytoplankton and cyanobacterial biomasses. This observation may be a consequence of lower amounts of nutrients being discharged into waterbodies from highly forested hydrological zone than from urban and/or agricultural areas. Our results illustrate a successful means of selecting representative waterbodies to conduct a regional assessment of cyanobacteria distribution using accessible GIS analyses.

Keywords: Cyanobacteria; Geographic information system; Lakes and reservoirs; Land use; Sampling strategy

1. Introduction

Cyanobacteria are ubiquitous components of phytoplankton communities in continental surface waters (Chorus and Bartram, 1999). They are frequently involved in bloom events, which lead to ecological disruption. They may also involve health risks for human and animal populations if the blooming species are potential toxin producers (Huisman et al., 2005). Due to their ecological and/or health importance, considerable efforts have been made to study the dynamics of these microorganisms (e.g. Rolland et al., 2005). There has been less investigation of their spatial distribution and the factors that control it, because, depending on the considered scale, it would require very large and expensive sampling. This is particularly true when sampling campaigns based on ecological or health considerations concern a whole region or an even wider geographical area including many lakes and reservoirs.

However, sampling on a large geographical scale is a very consistent way to address important scientific and/or management issues, such as i) describing and identifying the species that are found within a given geographical area, ii) evaluating species cosmopolitanism or endemism by assessing their spatial distribution and abundance, iii) estimating the degree of contamination of lakes and reservoirs by potentially toxic species within a given geographical area, which contributes to health risk assessment; and provides valuable information for interregional comparisons, and iv) identifying which environmental variables are related to changes in the distribution of the species of interest and could contribute to developing ecological models and/or guidelines for waterbody management.

The aim of a sampling strategy is to provide an unbiased sample of the target population (Hulbert, 1984). Ecologists studying macroorganisms have produced a large body of literature devoted to sampling strategy (Elliott, 1971; Pielou, 1975; Southwood, 1978; Frontier, 1983; Collwel and Coddington, 1995; Magurran, 1995). Some

sampling strategies make it possible to avoid selection bias, and these include simple random, stratified cluster and systematic sampling methods.

A review of the literature on phytoplankton/cyanobacterial mesoscale surveys of lakes and reservoirs (e.g. Duarte and Canfield, 1992; Willén and Mattsson, 1997; Reynolds and Peterson, 2000; Chorus, 2001; Cook et al., 2004; Willame et al., 2005) reveals that the issue of defining an unbiased sample population is either not considered at all or, if it is, not explicitly stated. Morris et al. (2002) have also pointed out that a large number of papers dedicated to the study of microbial diversity in general do not report the sampling strategy, especially in the field of aquatic microbiology. This situation may have resulted from (i) underestimating the critical importance of defining and using an unbiased sampling strategy, (ii) available data being obtained from previous surveys in which only the convenience of sampling was considered, (iii) difficulty of accessing some of the sites included in the sampling plan, (iv) limiting the amount of sampling performed in the study due to the time-consuming and/or costly methodologies used to measure the microbiological variables, and (v) a lack of basic data concerning the limnological or environmental characteristics of the lakes and reservoirs used to define the sampling plan.

However, with regard to the last two points, recent methodological developments can be used to improve our ability to define and apply unbiased sampling strategies to large-scale cyanobacterial distribution studies in continental waterbodies. Real-time methods based on *in vivo* fluorescence of photosynthetic pigments are now available for the study of microbial communities. They provide information about the total phytoplankton biomass, and *in vivo* fluorescence make possible to distinguish between the main taxonomic groups (Gregor et al., 2007). Despite the low level of taxonomic information, it provides information within seconds, and so greatly reduces the costs of analysis. Recent developments in remote sensing and GIS

(Geographic Information System) technologies have created new ways of describing lakes and reservoirs, as well as their surrounding environment. Indeed, it is now widely accepted that the nature of land use (Young et al., 1996; Tong and Chen, 2002; Mehaffey et al., 2005), as well as the geomorphologic and geological characteristics (Kratz et al., 1997; Willame et al., 2005) of the watershed, impacts on water quality in rivers and lakes and, as a consequence, on biological communities including the phytoplankton. Valuable data are now readily available in developed countries, however regionalization frameworks applied to sampling surveys (for a review, see Dentenbeck et al., 2005) are not widely used for the monitoring of ecological and natural resources. These characteristics are very useful for evaluating the nature and level of pressure faced by a given waterbody, which will affect water quality and, as a consequence, shape its phytoplankton communities (Knoll et al., 2003; Jones et al., 2004; Galbraith and Burns, 2007).

In this study, we used GIS data and database management tools to design a representative sampling strategy for lakes and reservoirs at the scale of the Ile-de-France (IDF) region based on hydrographical zones attributes. The selected waterbodies were sampled using a fluorescence-based field probe to estimate phytoplankton biomass and the relative abundance of cyanobacteria. The aims of this study were (i) to propose and test a sampling design for estimating cyanobacterial populations at a mesoscale, (ii) to evaluate the feasibility of using such a design to obtain a snapshot of the situation within a short time span, (iii) to find out whether the heterogeneity and diversity of the sampled sites are representative of the target population, and finally, (iv) to estimate the level of cyanobacterial biomasses in lakes and reservoirs, and the contribution of cyanobacteria to the overall phytoplanktonic biomass within the IDF region.

The IDF region constitutes a good “model” for studying the impact of human pressures on waterbodies, because it lies within a

single, first-order hydro-ecoregion as defined by Wasson et al. (2002) and thus, its geology is considered to be homogeneous at all the sampling sites. Furthermore, the IDF region displays a diversity of anthropogenic pressure in terms of its nature and intensity, and it includes a wide range of lakes and reservoirs. This regional scale also links many issues regarding water surface bodies and human societies: from human pressures impacts on ecosystems, to attempts by the local authorities to preserve them and in turn the societal and economic impact of the sanitary/ecological status of lakes and reservoirs. As a consequence, IDF provides a good mesoscale example for evaluating the risk associated to cyanobacterial development towards human populations, and its relationship to the nature and intensity of human pressures.

2. Material and Methods

2.1 Sampling area

The area investigated in this study is the Ile-de-France (IDF) region, an area covering 12011 km² in north-central France, in the center of the Paris basin (Figure 1a). With its eight administrative departments, Ile-de-France is home to 19% of the population of France (Marchand, 2007). It is a fertile sedimentary plateau where the Marne and Oise rivers join the Seine (Figure 1a). The region contains many large industrial towns and residential suburbs, and some agricultural production, mostly sugar beet and wheat. Despite its large population, half of the total area of the region is still used for agricultural purposes (Bertaud et al., 2004). Forested areas cover over 23% (i.e. 278 000 ha) of the region and are mainly found in locations unsuitable for intensive farming (IFN, 2007).

2.2 Target population and sample size

The study area is composed of 990 surface water units (lakes, reservoirs and ponds; Figure 1a) as defined by the RNDE (Réseau National des Données sur l'Eau, France), a governmental organization that centralizes data about French water resources. These

surface water units were identified using the Carthage 3.0[®] GIS database (IGN, Paris, France).

From these 990 surface waterbodies (Figure 1), we chose to study only those that cover an area of more than 5 hectares ($n = 248$), as they are the most used by human populations (generally open to the public for water sports, fishing etc.). In addition, we only took into account those found in hydrographical zones which lie entirely within the IDF region ($> 95\%$, $n = 50$; Figure 1b), which amounts to 188 lakes and reservoirs (Figure 1a). The sampling effort was set to 50 lakes, reservoirs or ponds.

2.4 Sampling design

We used a stratified simple random sampling design, a method considered as efficient to obtain representative samples from heterogeneous environments, while simultaneously minimizing the time and cost of conducting surveys (e.g. Frontier, 1983). The steps involved in constructing the sampling design are summarized in Figure 2. After building the strata (see below), the 50 sites to be sampled were allocated among the different defined strata in proportion to the number of targeted waterbodies within each stratum. Sites within each stratum were randomly selected.

Hydrographical zones as the spatial unit of the stratification process

The first step was to choose the spatial unit to be used to build the strata. Ideally, the true watersheds of waterbodies are the finest and most appropriate spatial units on which strata can be built. However, data at the watershed scale were not available. These resource constraints led us to use the hydrographical zone (HZ, Figure 1b) as the reference spatial unit. The French hydrographical codification, in accordance with European directive 2000/60/DCE (23/10/2000), defines the hydrographical zone as elementary zones of which boundaries are based on topographical watersheds. They correspond to the finest

watersheds sub-units delineation framework available in GIS databases. The hydrographical zones are defined in the same way as the 8-digit hydrologic units (HUC-8s) used by the United State Geological Survey (Seaber et al., 1987). In their study, Cheruvelil et al. (2008) showed that HUC-8s was the clustering framework that maximized the information among lakes groups.

The hydrographical zone scale provides a hydrological functional delineation for which both topographical and environmental data were available over the whole studied area. Hydrographical zones boundaries are available in the Carthage 3.0[®] GIS database.

Hydrographical zone characterization

To be able to take into account the different potential anthropogenic pressures that waterbodies can experience, we described the typology of hydrographical zones through indices (Table 1) based on land use (cropland (LU_c), urban impervious (LU_i) and non-agricultural pervious (LU_p)), the relative size of the hydrographical zones (I_{rs}), drainage index (I_d) and altitude index (I_s). Some of the other indices that we could have used were eliminated from the analysis due to collinearity with other variables (e.g. impervious surface vs. population density, $R^2 = 0.96$, $p < 10^{-5}$). All indices were designed to range from 0 to 1 and would therefore have the same weight in the similarity analysis (see below).

These environmental indices were defined using information available in the GIS databases. GIS data were analyzed using MapInfo 8.0 Professional[®] (MapInfo Inc., Troy, USA). Land use indices were calculated using data from the 2005 version of the DensiMos[®] GIS database (IAURIF, Paris, France), which includes data from the last completed population inventory (1999), and data about land use (2003). Data relative to relief were handled using Vertical Mapper 3.1[®] (TETRAD Computer Applications Inc., Bellingham, USA), an add-on to MapInfo.

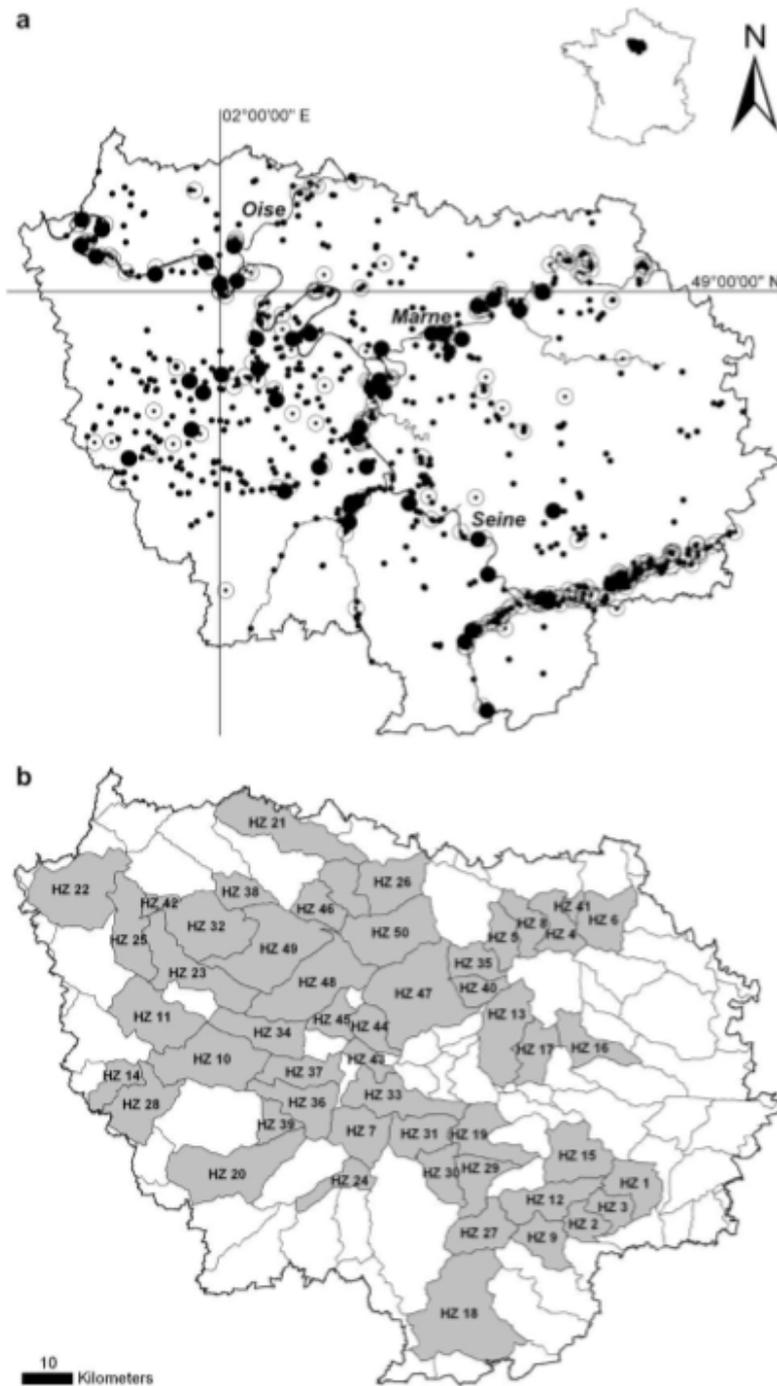


Fig. 1. (a) Map of the Ile-de-France region (France) showing the location of the 980 lakes, reservoirs and ponds included in the Carthage 3.0[®] GIS database (dots; n = 980), waterbodies > 5 ha (open circles; n = 248) and sampled waterbodies (black circles; n = 50). (b) Hydrographical zones (HZ) in the IDF region (n = 134). The hydrographical zones included in the study are shown in gray (surface area within the IDF > 95%, and containing waterbodies covering over 5 ha; n = 50).

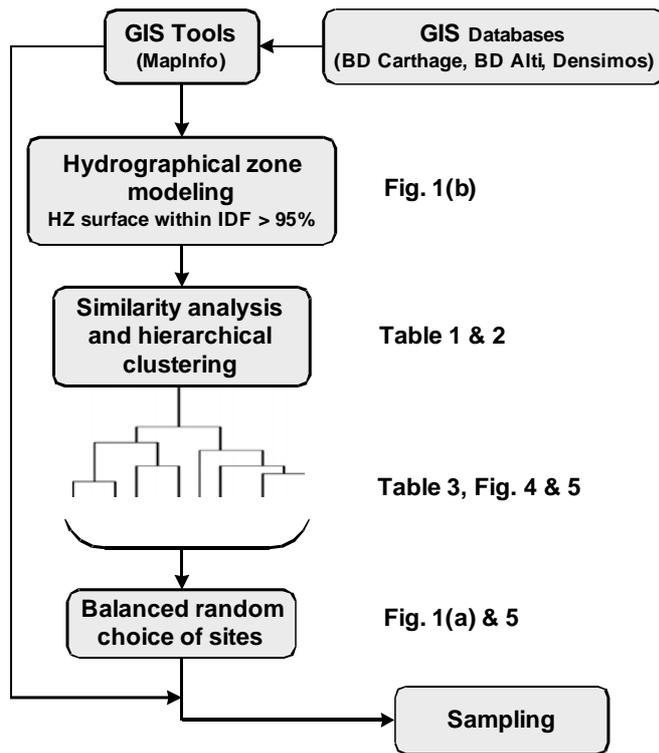


Fig. 2. Integrated sampling strategy based on hydrographical zones environmental characteristics.

Index		
<i>Morphology</i>		
Relative surface index	I_{rs}	$(S_{HZ}/\sum S_{HZ})/(S_{HZ}/\sum S_{HZ})_{max}$
Relative altitude index	I_{alt}	$(Alt_{av})/(Alt_{av})_{max}$
<i>Land-use</i>		
Impervious surface index	LU_i	$(S_i/S_{HZ})/(S_i/S_{HZ})_{max}$
Agricultural surface index	LU_c	$(S_a/S_{HZ})/(S_a/S_{HZ})_{max}$
Non-agricultural pervious surface index	LU_p	$(S_p/S_{HZ})/(S_p/S_{HZ})_{max}$
<i>Hydrology</i>		
Drainage index	I_d	$(L_{riv}/S_{HZ})/(L_{riv}/S_{HZ})_{max}$

S_{HZ} : surface area of the hydrographical zone; S_a : agricultural surface area (%); S_i : impervious surface area (%); S_p : pervious non-agricultural surface area (%), which corresponds to grassland and forest; L_{riv} : total length (km) of river streams within a hydrographical zone per square kilometers; Alt_{av} : average altitude of a hydrographical zone.

Table 1. Environmental variables used to define the types of hydrographical zones in the Ile-de France region.

The drainage index was calculated using surface water data available in the Carthage 3.0[®] GIS database.

Typology of the hydrographical zones of IDF
Similarity analysis (Gower similarity coefficient; Legendre and Legendre, 1998) based on the defined environmental indices was then applied to build the clusters (i.e. strata) of hydrographical zones (Ward's

minimum variance method, Legendre and Legendre, 1998). Similarity analysis and hierarchical clustering were performed using the R package (Casgrain & Legendre, 2004).

2.5 Field sampling and analysis of the samples

Three stations were sampled within each site to estimate spatial heterogeneity. Sampling stations were located using MapInfo[®] to

maximize the distance between stations according to the morphology of the waterbodies. The stations were then located in the field using a portable GPS apparatus (GPS12, Garmin, Olathe, USA). All samples were taken 0.5 m below the water surface.

One and a half liter samples were placed in plastic bottles, and kept refrigerated and protected from light sources until being filtered through 47 mm Whatman GF/F glass fiber filters in the laboratory. Filters were frozen at -24 °C. Chlorophyll *a* (Chl*a*) was extracted using 7.5 ml of methanol/water (90/10 v/v) for spectrophotometric quantification, according to Talling & Driver (1963). Extracts were analyzed using a Cary 50 Scan spectrophotometer (Varian Inc., Palo Alto, USA).

Five liter samples were placed in dark plastic bottles, and used to perform phytoplankton biomass and composition estimation using a FluoroProbe II device (bbe-Moldaenke, Kiel, Germany). This probe contains five diodes emitting light at 450, 525, 570, 590 and 610 nm for phytoplankton pigment excitation. A 370 nm diode is also included for the excitation of dissolved organic substances. The fluorescence emitted by these molecules was then subtracted to obtain the fluorescence signal due to the phytoplankton alone. This device is able to distinguish between four main phytoplanktonic groups (i.e. diatoms and dinoflagellates, blue-green algae, green algae and cryptophytes) by comparing the excitation spectrum obtained from natural samples to standard curves stored in the probe. The relative amount of each class was then calculated, and expressed in terms of the equivalent amount of Chl*a* per liter of water ($\mu\text{g equiv.Chl}a.L^{-1}$). For a detailed description of the apparatus, see Beutler et al. (2002). The samples were rigorously protected from light sources for 30 minutes before taking the measurements to avoid attenuation of the fluorescence signal due to photochemical quenching. Measurements were performed every second for two minutes, yielding 120 data points per sample. The total Chl*a* values obtained with the probe were compared to those obtained from

spectrophotometric analysis. A type-II linear regression was fitted to the data to obtain the conversion factor to be applied to the probe data. A type-II linear regression was used as the two variables are both random and measured with error (Legendre and Legendre, 1998). The equation computed was $Y = 1.816 * X - 2.080$ ($n = 147$; $R^2 = 0.868$; $p < 10^{-4}$; data not shown), and the FluoroProbe data were corrected accordingly.

Sampling was carried out in the summer of 2006 (from 28 July to 13 August). The sampling survey was as short as possible (about two weeks) to provide a 'snapshot' of the phytoplankton communities at the 50 sampled sites, and to reduce variability due to differences in meteorological conditions.

2.6 Data analysis

The trophic state of the sampled waterbodies was defined following Ryding and Rast (1989) guidelines, using the maximum total chlorophyll *a* values.

Contingency analyses using Wilks chi-square corrected by Williams (1976), as described by Legendre and Legendre (1998), were carried out to find out whether there was a significant association between global land use categories (urban, periurban, rural) and the trophic level category of the sampled lakes and reservoirs within the IDF. We decided to use Wilks chi-square and not Pearson chi-square because the former is more conservative for small numbers of observations. Contingency analyses were performed using MS Excel 2003 (Microsoft Corporation, Redmond, USA). The other statistical analyses were performed using GraphPad Prism version 4.0 for Windows (GraphPad Software, San Diego, USA).

3. Results

3.1 Hydrographical zones characteristics

The 50 hydrographical zones investigated in this study displayed a gradient of land use and land cover types (Table 2). Figures 3 and 4 show three contrasting regions in terms of human use. Eight hydrographical zones

contained a high proportion of impervious surface (over 50%; Figure 3) corresponding to an area (Figure 4 c) including Paris and its inner suburban area at the center of the IDF region. Twenty-two hydrographical zones were covered by over 50% agricultural or 50% non-agricultural pervious surface (Figure 3), and formed a rural belt in the outer part of the IDF region (Figure 4 d). The remaining 20 hydrographical zones displayed intermediate characteristics, and were located between the urban and rural areas. This division into three belts (urban, periurban and rural) can easily be observed by considering urban cover alone (see Figure 4 c).

The hydrographical zones covered surface areas between 18 and 326 km² (Table 2). Within the IDF landscape they were distributed between low altitude and highly drained basins located around the main rivers (Seine, Marne and Oise) and higher altitude areas on top of the sedimentary plateau (Table 2, Figure 4 b and f).

The characteristics defined above show that the IDF region is exposed to a gradient of potential human use linked to human activities and diverse geomorphologic environments.

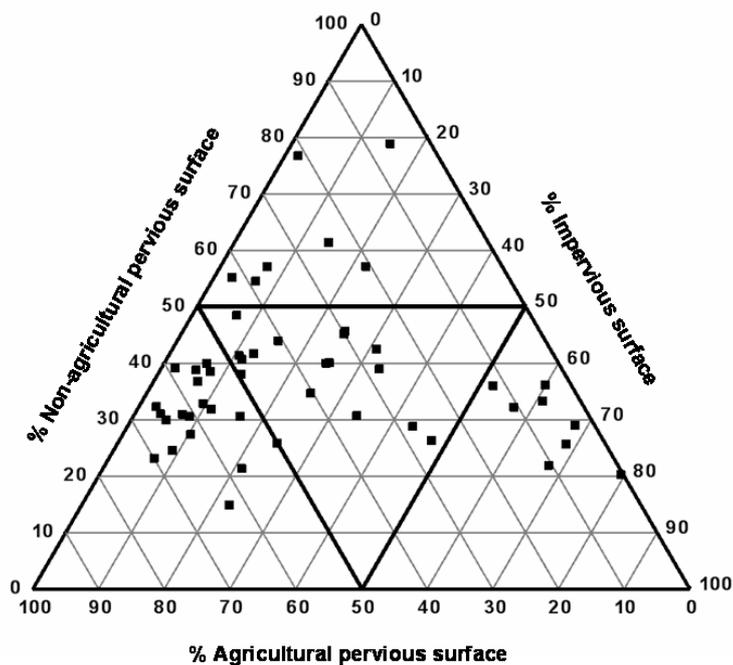


Fig. 3. Ternary diagram showing the proportions of non-agricultural pervious surface (forest and non-agricultural grassland), agricultural (cropland and grassland) surface, and impervious (urban) surface within the studied hydrographical zones.

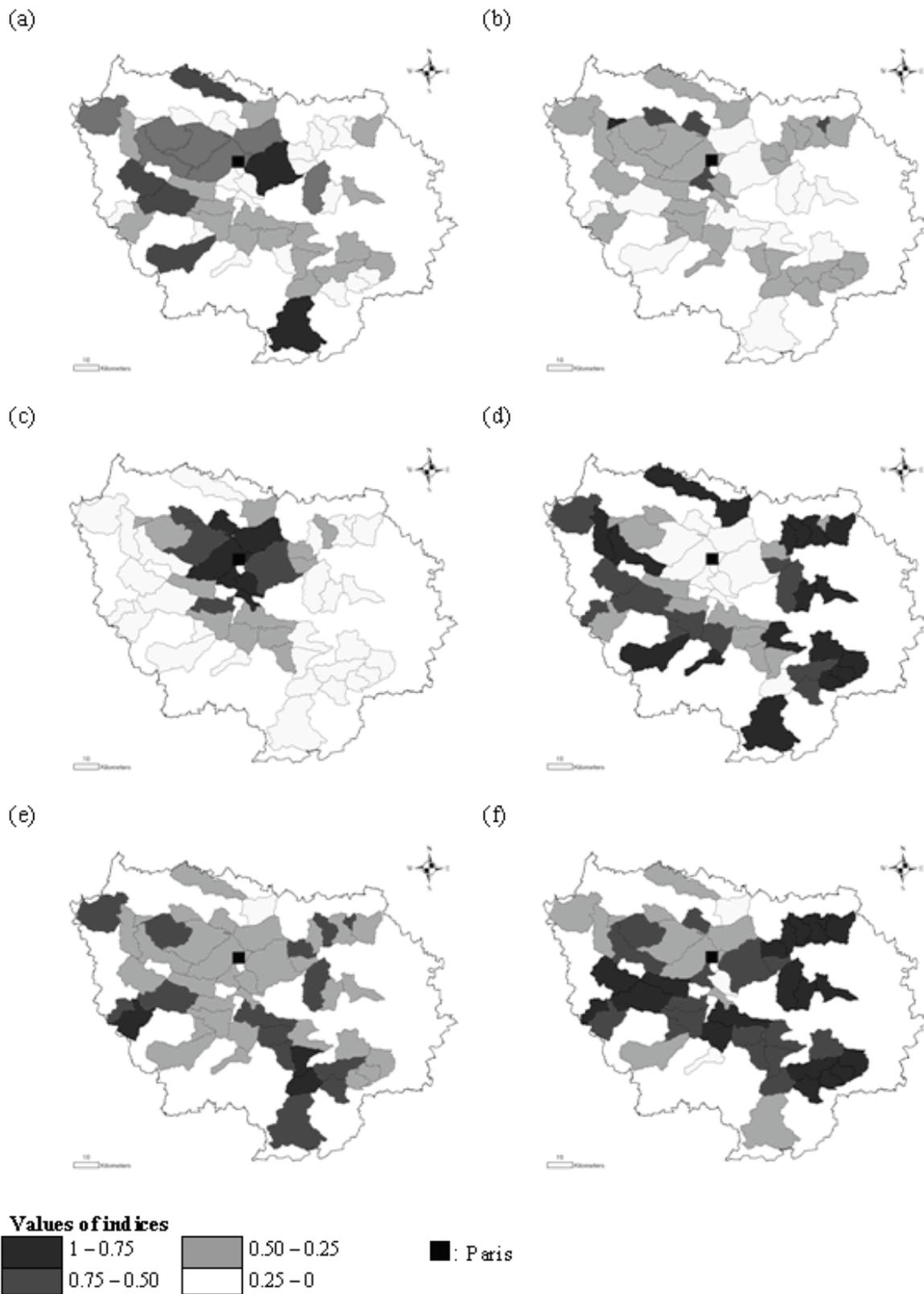


Fig. 4. Characteristics of the hydrographical zones (a) Relative surface index; (b) Relative altitude index; (c) Impervious surface index; (d) Agricultural surface index; (e) Non-agricultural pervious surface index; (f) Drainage index.

Table 2. Summary statistics for geomorphology, hydrographical characteristics, land cover and population density in the Ile-de-France hydrographical zones (n = 50).

	Median	Mean	Min	1 st quartile	2 nd quartile	Max
<i>Hydro-geological characteristics</i>						
Surface area (km ²)	100	116	18	65	158	326
Average altitude (m)	92	96	51	76	109	161
Drainage index (m.km ⁻²)	601	606	206	471	801	952
<i>Land-use and population density</i>						
Urban (%)	14.6	23.1	2.0	7.1	29.6	79.4
Agriculture (%)	40.7	37.1	0.1	23.6	55.2	69.6
Non-agricultural grassland (%)	12.5	12.8	0.3	8.4	17.3	28.3
Forest (%)	22.2	24.3	1.6	15.5	32.8	70.0
Water (%)	1.8	2.1	0.1	0.6	3.2	8.4
Population density (pop.km ⁻²)	464	1426	31	207	1096	13340

3.2 Typology of hydrological zones and site selection

Based on the 4 environmental and 2 human use variables identified in Table 1, similarity analysis combined with hierarchical clustering made it possible to define hydrographical zone groups (see Figure 5 and Table 3). We retained a similarity level allowing obtaining 10 hydrographical zones groups. This level was chosen because (i) it

gave homogeneous clusters in terms of environmental pressure and hydromorphological characteristics; (ii) it allowed maximizing the probability of sampling sites encountering unusual environmental contexts and (iii) it produced clusters containing a sufficient number of lakes to sample at least one individual per cluster using the method described below.

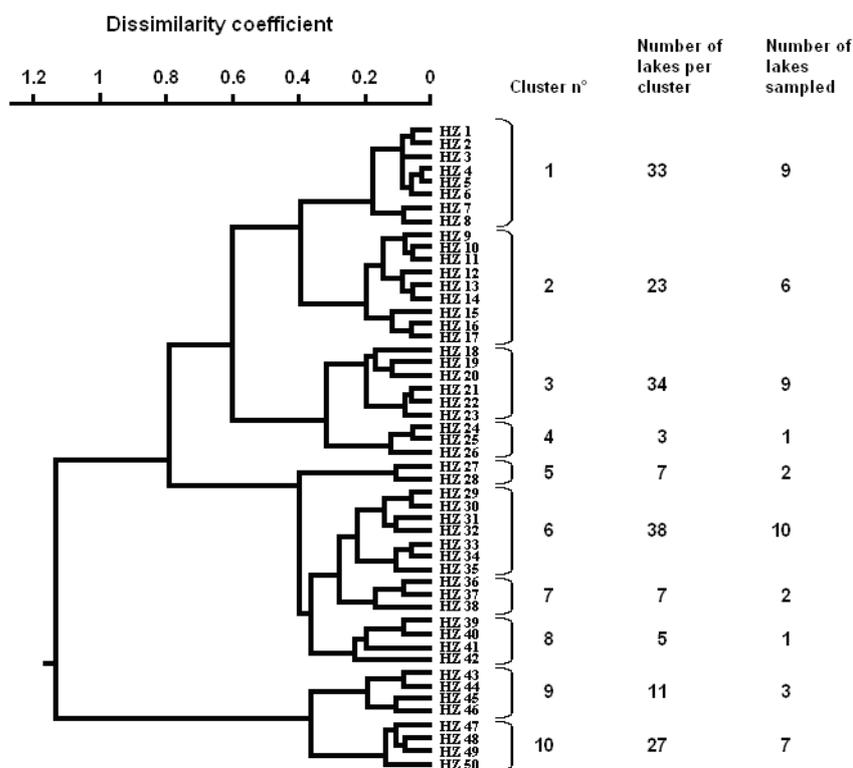


Fig. 5. Hierarchical clustering based on the similarity analysis (Gower, S15) of the characteristics of the hydrographical zones.

Within each cluster, a number of lakes and reservoirs were randomly selected proportionally to the number of potential sampling sites per cluster (Figure 5). From

the 50 sites originally selected, two could not be accessed. These sites were replaced by two new randomly-selected sites within the corresponding clusters.

Table 3. Average index values for the ten clusters identified by similarity analysis.

	I_{rs}	I_{alt}	LU_i	LU_c	LU_p	I_d
Cluster 1	0.2381	0.3422	0.1376	0.8644	0.3561	0.9124
Cluster 2	0.3909	0.1967	0.0990	0.6827	0.5520	0.8188
Cluster 3	0.6365	0.2629	0.1176	0.8016	0.4303	0.4666
Cluster 4	0.3160	0.3807	0.1822	0.8394	0.3337	0.2441
Cluster 5	0.2745	0.2788	0.1083	0.1767	0.9819	0.6246
Cluster 6	0.3252	0.2640	0.3255	0.3613	0.6069	0.6495
Cluster 7	0.2720	0.4191	0.5312	0.4108	0.3606	0.5754
Cluster 8	0.1012	0.6021	0.2256	0.5242	0.5643	0.7136
Cluster 9	0.1656	0.4101	0.8600	0.0302	0.3658	0.4324
Cluster 10	0.6887	0.2303	0.7854	0.0996	0.3793	0.4284

I_{rs} : Relative surface index; I_{alt} : Relative altitude index; LU_i : Impervious surface index; LU_c : Agricultural surface index; LU_p : Non-agricultural pervious surface index; I_d : Drainage index.

3.3 Limnological characteristics

The 50 waterbodies selected through the sample strategy are somehow artificial. Most originated as sand and gravel quarries worked between the 1940s and 1980s, or resulted from peat extraction in the mid 19th century. The older waterbodies are reservoirs built in the 17th and 18th centuries to provide a reliable water supply for Versailles Castle and the city of Paris. As a consequence of their anthropogenic origin, a very unusual characteristic of these waterbodies is that nearly half of them ($n = 24$) are not connected to the hydrological network (Table 4). Most of the waterbodies can be classified as shallow (depth < 5 m, $n = 47$), and range from 5 ha to over 120 ha (Table 4).

On the basis of the chlorophyll *a* data (Table 4), the selected waterbodies showed a wide

range of phytoplanktonic biomass (mean $Chla < 1 - 357 \mu g \text{ equiv. } Chla.L^{-1}$). According to the eutrophication levels defined by Ryding and Rast (1989), 36 of the waterbodies can be considered to be eutrophic ($n = 15$) to hypereutrophic ($n = 21$). Waterbodies joined to the hydrographical network ($n = 26$) and the isolated waterbodies ($n = 24$) had significantly different chlorophyll *a* concentrations (Mann-Whitney U-test, $p < 0.0001$), with the isolated waterbodies displaying lower mean chlorophyll *a* concentrations. No relationship was observed between surface, depth or volume of waterbodies and $Chla$ for any of the tested population (entire sampled population, isolated or connected waterbodies).

Table 4. Location, morphological characteristics of the 50 sampled sites and phytoplanktonic biomass observed in august 2006 (FluroroProbe data). (d.l.: detection limit).

Name	Cluster	Localisation	Linked to rivers and streams	Surface area (ha)	Mean depth (m)	Secchi depth (m)	Total Chl <i>a</i> ($\mu\text{g}\cdot\text{L}^{-1}$)	Cyanob. Chl <i>a</i> ($\mu\text{g}\cdot\text{L}^{-1}$)	Cyanob. biomass (%)
1- Meaux Park Pond	1	48°56' N 02°54' E	-	6.8	1.1	0.58	32.4	17.6	54.4
2- Précy-sur-Marne Pond	1	48°55' N 02°45' E	-	26.6	3.2	0.98	7.0	2.7	39.0
3- Isles-les Villenoy Pond	1	48°54' N 02°50' E	-	41.6	2.6	1.35	10.1	2.4	23.5
4- Fontenay/Vic Pond	1	48°33' N 02°23' E	+	9.8	1.2	0.27	357.3	278.7	78.0
5- Vert-le-Petit Pond	1	48°33' N 02°22' E	-	5.0	2.0	0.50	123.1	57.5	46.7
6- Fleuri Pond	1	48°32' N 02°22' E	+	8.4	1.3	0.62	67.6	9.9	14.6
7- Courcouronnes Pond	1	48°37' N 02°25' E	+	6.5	1.4	0.32	123.5	92.1	74.6
8- Delomez Pond	1	48°24' N 03°07' E	-	6.7	4.2	0.83	9.6	0.6	6.3
9- Leclerc Pond	1	48°24' N 03°05' E	-	5.4	2.8	0.82	6.1	0.6	10.1
10- Maurepas Reservoir	2	48°46' N 01°55' E	+	7.6	3.4	0.60	138.7	24.7	17.8
11- Noés Reservoir	2	48°45' N 01°58' E	+	24.0	0.7	0.28	73.0	17.1	22.4
12- Vaux-de-Cemay Pond	2	48°41' N 01°56' E	+	6.5	1.4	0.65	101.6	34.1	33.6
13- Grande Paroisse Pond	2	48°22' N 02°54' E	+	52.7	3.7	0.37	191.5	185.8	97.0
14- Grand Marais Pond	2	48°22' N 02°57' E	-	16.5	2.4	1.08	27.1	14.6	53.7
15- Villefermoy Reservoir	2	48°32' N 02°56' E	+	21.7	2.1	0.38	54.8	19.7	35.9
16- Freneuse Cove	3	49°04' N 01°35' E	+	14.0	2.4	0.47	202.3	31.6	15.6
17- Grand Lavacourt Pond	3	49°03' N 01°41' E	-	111.5	4.3	0.80	25.3	12.8	50.6
18- Ilon Cove	3	49°01' N 01°37' E	+	7.4	3.8	0.55	84.5	19.8	23.5
19- ASM Club Cove	3	49°00' N 01°40' E	+	9.2	2.5	0.45	181.4	39.2	21.6
20- Bout du Monde Pond	3	48°58' N 01°50' E	-	27.6	1.2	0.87	3.1	0.5	15.0
21- Port-Sud Pond	3	48°34' N 02°11' E	+	6.5	1.7	0.48	118.8	66.1	55.6
22- Clarette Pond	3	48°19' N 02°42' E	-	13.9	1.4	0.30	129.7	12.2	9.4
23- Fromonville Pond	3	48°17' N 02°41' E	-	6.5	3.3	0.93	8.5	1.1	12.9
24- Souppes/Loing Pond	3	48°10' N 02°45' E	+	21.1	2.9	0.90	39.7	5.1	12.8
25- Saint-Blaise Pond	4	48°31' N 02°22' E	+	5.2	1.4	0.67	68.6	9.8	14.2
26- Gazeran Pond	5	48°38' N 01°46' E	+	8.2	0.7	0.28	45.7	9.2	20.0
27- Samoreau Pond	5	48°25' N 02°45' E	-	10.2	2.4	1.40	2.4	0.7	27.7
28- Grosse Pierre Pond	6	49°00' N 01°58' E	-	46.2	4.3	1.03	34.2	15.4	45.1
29- Gaule Achéroise Pond	6	48°58' N 02°03' E	-	5.3	6.2	1.72	7.8	3.3	41.7
30- Triel Pond	6	48°57' N 02°00' E	-	34.1	3.9	n.d.	61.4	56.1	91.4
31- Jabline-Anet Lake	6	48°54' N 02°43' E	-	77.1	7.0	4.10	6.5	0.3	4.7
32- Saclay Reservoir	6	48°44' N 02°10' E	+	30.2	1.5	0.52	77.0	12.3	15.9
33- La Veyssière Pond	6	48°41' N 02°24' E	+	27.1	1.6	0.47	143.5	64.6	45.0
34- Noues de Seine Pond	6	48°40' N 02°23' E	+	29.1	3.4	0.77	49.9	10.7	21.5
35- Seine Port Pond	6	48°33' N 02°32' E	-	5.9	2.9	0.49	176.5	139.5	79.0
36- Saint-Quentin Reservoir	6	48°47' N 02°01' E	+	120.1	1.6	0.32	103.7	42.0	40.5
37- Bois-le-Roi pond	6	48°29' N 02°43' E	+	8.3	2.6	0.67	23.3	13.5	58.0
38- Les Galets Pond	7	49°01' N 02°03' E	-	10.6	2.2	1.45	15.0	7.7	51.4
39- Epinoche Pond	7	48°37' N 02°17' E	+	5.7	2.4	0.58	110.4	44.7	40.5
40- Loy Pond	8	48°51' N 02°41' E	+	6.9	1.3	0.37	225.3	99.1	44.0
41- Créteil Lake	9	48°46' N 02°27' E	+	40.8	4.5	0.73	20.7	1.8	8.8
42- Choisy Pond	9	48°46' N 02°25' E	-	21.6	2.6	0.83	39.6	5.2	13.2
43- Plage Bleue Pond	9	48°45' N 02°28' E	-	9.3	3.1	0.72	27.4	5.9	21.5
44- Lake Inférieur	10	48°51' N 02°16' E	+	9.5	1.2	0.30	255.3	132.8	52.0
45- Les Pâtis Pond	10	48°52' N 02°36' E	-	10.3	2.6	1.33	35.8	10.8	30.2
46- UCPA Centre Pond	10	48°52' N 02°37' E	-	92.0	5.3	4.22	1.6	0.4	25.5
47- Saint-Cucufa Pond	10	48°51' N 02°10' E	+	5.4	1.8	0.77	85.0	3.6	4.3
48- UTE Louveciennes	10	48°51' N 02°06' E	-	7.5	2.5	n.a.	< d.l.	< d.l.	-
49- Lake Minimes	10	48°50' N 02°27' E	+	5.8	1.2	0.40	194.4	61.7	31.7
50- Swiss Pond	10	48°48' N 02°07' E	-	14.4	1.9	0.83	32.5	14.3	44.1

3.4 Cyanobacterial biomass

The cyanobacterial biomass ranged from less than the detection limit ($1\mu\text{g equiv.Chla.L}^{-1}$) to over $278\mu\text{g equiv.Chla.L}^{-1}$ (Table 4). Cyanobacteria dominated the phytoplankton communities in 12 waterbodies, and contributed over 90% of the phytoplankton biomass in two waterbodies (13-Grande Paroisse Pond and 30-Triel Pond). Highly dominant cyanobacterial levels were more frequent in the most eutrophic waterbodies. A significant positive relationship was found between the total phytoplankton biomass and (i) the cyanobacterial biomass ($r = 0.849$, $p < 0.0001$, $n = 49$) and (ii) the percentage of chlorophyll *a* relative to cyanobacteria ($r = 0.393$, $p < 0.01$, $n = 49$). The cyanobacterial biomass was significantly lower in isolated waterbodies than in waterbodies connected to the hydrographical network (Mann-Whitney U-test, $p < 0.001$).

3.5 Phytoplankton distribution and land uses

We used contingency analysis to check whether our sampling methodology revealed a significant relation between the nature of land use and the trophic status of waterbodies in the IDF region. The analysis was performed using four trophic status classes (oligotrophic ($n = 8$), mesotrophic ($n = 7$), eutrophic ($n = 14$) and hypereutrophic ($n = 21$)), as defined by Ryding and Rast (1989), and three global land use groups in this study (urban ($n = 10$), periurban ($n = 25$) and rural ($n = 15$), as described above) as shown in Figure 6. Contingency analysis did not reveal any significant relationship between trophic status and global land use groups ($X_{w\text{ corr}} = 11.768$, $p = 0.067$, $df = 6$, $n = 50$).

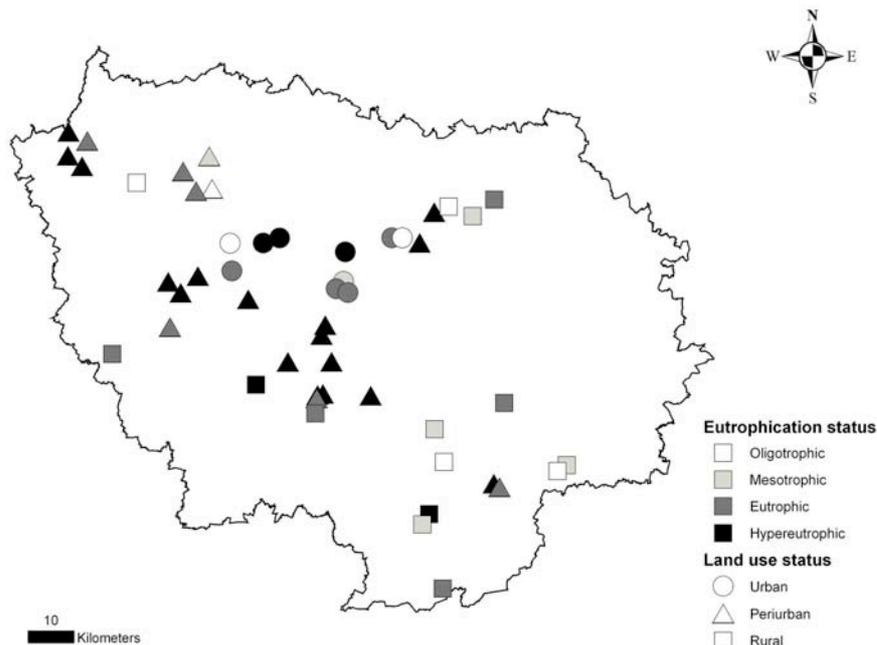


Fig. 6. Eutrophication status (grey scale) and surroundings land use status (symbols shape) of the 50 surface waterbodies sampled in IDF.

We then investigated this relationship further by considering separately the waterbodies linked to the hydrographical network and the isolated ones. Contingency analysis showed a significant link between trophic status and global land use categories for waterbodies

connected to the hydrographical network ($X_{w\text{ corr}} = 10.121$, $p = 0.038$, $df = 4$, $n = 26$), but not for the isolated ones ($X_{w\text{ corr}} = 10.562$, $p = 0.103$, $df = 6$, $n = 24$). To further investigate the significant link observed between land use categories and the trophic status of

waterbodies joined to the hydrographical network, we examined the relationship between both total phytoplankton and cyanobacterial biomass and the percentage of hydrographical zones surface covered by the 5 land use types defined in Table 2. A significant negative correlation (non-parametric spearman correlation) was found between the percentage of forest cover and both total phytoplankton ($r_s = -0.589$, $p < 0.01$, $n = 25$; Figure 7a) and cyanobacterial biomass ($r_s = -0.497$, $p < 0.05$, $n = 25$; Figure 7b). However, the large scattering of data

points precludes to satisfactorily fitting either a linear (total phytoplankton: $R^2 = 0.330$; cyanobacterial biomass $R^2 = 0.241$) or non-linear (e.g. decreasing exponential model; total phytoplankton: $R^2 = 0.345$; cyanobacterial biomass: $R^2 = 0.330$) model to our observations. Computing the Akaike's Information Criteria (AIC) between the two alternative non-nested models showed a low probability to correctly prefer either of the tested models (total phytoplankton: $p = 0.487$; cyanobacterial biomass: $p = 0.102$).

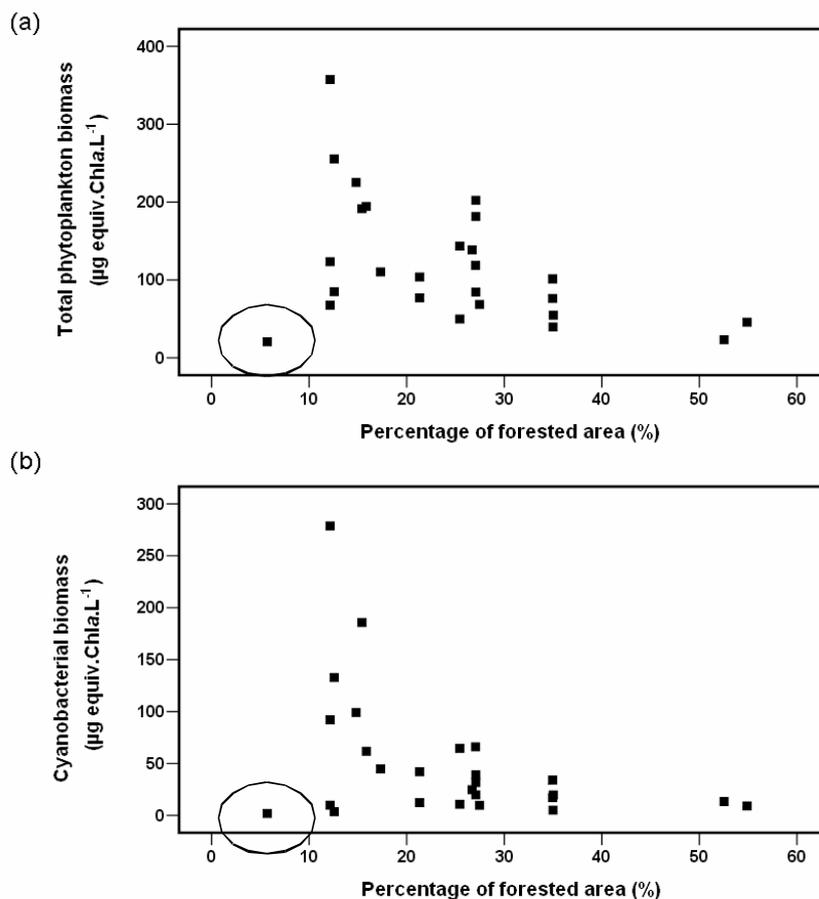


Fig. 7. Relationship between (a) total phytoplankton biomass ($\mu\text{g equiv.Chla.L}^{-1}$) and (b) cyanobacterial biomass ($\mu\text{g equiv.Chla.L}^{-1}$) against the proportion of forested surfaces (%) in the lakes and reservoirs connected to the hydrographical network ($n = 26$). The dot within the ellipse was considered to be an outlier as the connection to the hydrographical network is only hypothetical through a permanently closed weir.

4. Discussion

4.1 On the use of landscape data in sampling strategy definition

Sampling strategies used to evaluate the occurrence and toxic potential of cyanobacteria at a national or regional scale have mainly been driven by the human use of the waterbodies concerned (Duarte and

Canfield, 1992; Cook et al., 2004; Jayatissa et al., 2006) or by selecting sites known to present recurrent cyanobacterial problems (Willén and Mattson, 1997; Reynolds and Petersen, 2000; Willame et al., 2005).

However, the regional scale has been shown to be very informative for studying the links between natural (geomorphology, geology, hydrology...) and human (urban and agricultural pressure, local management...) pressures and the water quality of both lentic (Goslar et al., 1999; Jones et al., 2004) and lotic (Arnold and Gibbons, 1996, Tong and Chen, 2002, Mehaffey et al., 2005, Burford et al., 2007) systems. For instance, the position of the waterbody in the landscape determines the relative importance of groundwater and precipitation input to a lake (Kratz et al., 1997), whereas the watershed size correlates positively with the amount of nutrients that get into hydrologic systems (Young et al., 1996).

On the basis of this previous knowledge, the regionalization framework designed in this study allowed us to define the 10 major categories of environmental conditions that may confront the waterbodies of IDF. The 10 categories were then used to define a stratified sampling strategy for IDF waterbodies. The regionalization framework can be refined using additional environmental variables or indexes (e.g. wetland and riparian land cover) when available and/or to better describe local environmental conditions.

Our sampling design (50 waterbodies out of the 980 lentic water systems in IDF were sampled) was made possible by using *in-situ* fluorescence probe technologies that permit a quick estimation of phytoplankton and cyanobacterial biomasses (expressed in $\mu\text{g equiv.Chla.L}^{-1}$).

4.2 Eutrophication and cyanobacteria issues in the IDF

The results obtained showed first that all trophic status categories ($< \text{d.l. to } 357 \mu\text{g equiv.Chla.L}^{-1}$) occurred in the waterbodies of the IDF region. However, 72% of the waterbodies sampled could be classified as

eutrophic habitats (42% were hypereutrophic). Overall, this high percentage of eutrophic waterbodies is likely to be related to the high level of anthropogenic pressures exerted within the studied region.

Cyanobacteria were detected in 49 out of the 50 waterbodies sampled, and dominant in 24% of them. This value is in the lower range of the data found for temperate regions (Olding et al., 2000: 22%; Wiedner et al., 2001: 38%; Duarte and Canfield, 1992: 49%). However, these studies do not state how representative their sampling strategy was, and the chosen sampled sites may not constitute an unbiased representation of the diversity of the waterbodies found in the studied area.

Our results clearly show that most of the waterbodies within the IDF region can be considered as eutrophic/hypereutrophic, which may be expected in a region exhibiting strong anthropogenic pressures. The eutrophication level depends significantly on whether these waterbodies are linked or not to the hydrographical network, the former being more eutrophic than the latter. Despite the expected higher water residence time, isolated waterbodies may be subjected to lower anthropogenic pressure. Two complementary processes may support this hypothesis. The Seine–Normandy Water Authorities have considerably reduced point sources of pollution from domestic and industrial sewers over the last 30 years, and non-point source pollution is now considered to be the main cause of ecological deterioration of the waterbodies in the Seine–Normandy basin, which includes the IDF region (Even et al., 2007). The waterbodies isolated from the hydrographical network may collect water from a smaller drainage basin and therefore may receive less non-point source pollution (Young et al., 1996). However, one has to bear in mind that the potential loading impact on a water quality is related not only to the drainage basin size but also to the volume of the waterbody.

4.3 Land uses impact on lentic waterbodies and the hydrographical zone scale

Considering waterbodies isolated from the hydrographical network, we were unable to show any significant relationships between water quality variables and environmental features obtained at the hydrographical zone scale. However, significant relationships between the nature of land use in hydrographical zones and the trophic status of waterbodies were observed for waterbodies connected to the hydrographical network. The total phytoplankton and cyanobacterial biomass within these waterbodies showed (i) a negative correlation with the percentage of the hydrographical zone covered by forested surfaces and, (ii) no significant relationship to the proportion of area covered either by urban or agricultural surfaces. This may result from a lack of independence between the proportions of these two land-use categories (when agricultural land cover decreases, urban land cover increases and vice versa), which both are non-point sources of nutrient inputs. In turn, this may explain why the highest frequency of highly eutrophic waterbodies is found within hydrographical systems showing intermediate land use, with both relatively widespread urban and cropland cover. Jones et al. (2004) did not take into account the reservoirs located in urban areas to pinpoint the role of cropland area on water quality in Missouri reservoirs. In mixed landscapes, the proportion of forest or natural cover within a given hydrographical zone may be a good candidate for estimating potential nutrient inputs in waterbodies.

Using hydrographical zones helped stratifying the sample population and ensuring the distribution of the sampled sites across a gradient of land-use. However, stronger or unrevealed relationships may have been found between environmental characteristics of the waterbodies surroundings and water quality using true

watersheds as the reference spatial scale, instead of hydrographical zones.

5. Conclusion

Our results allowed getting a comprehensive view of the diversity of environmental conditions that can occur in the IDF. It constitutes a preliminary step and further analyzes of the links between cyanobacteria distribution in waterbodies and environmental variables are required.

With the addition of the identification of potential toxic species and/or the determination of toxins concentrations, such a strategy would help to assess appropriately the sanitary risk linked to human uses of waterbodies within the IDF.

The improved availability of GIS data, data-handling and methodological capabilities allow one to implement such sampling strategies in a wide range of spatial scales and environmental questions. It could thus be a useful tool to maximize the information obtained from the sampling of waterbodies and required by the European directive 2000/60/DCE (23/10/2000) for the environmental assessment of waterbodies quality.

Acknowledgements

This work was funded by the ECODYN (INSU n°04CV131) and ANR SEST 07 Cyanotox research program. We would like to thank N. Escoffier and C. Yéprémian for field sampling, technical assistance and helpful discussions. This work was made possible by a MNHN fellowship (Legs Prévost, Conseil Général de Seine Saint-Denis) awarded to A. Catherine.

References

- Arnold, C.L. and Gibbons, C.J. (1996) Impervious surface coverage: the emergence of a key environmental indicator. *Journal of the American Planning Association* 62(2), 243-258.
- Bertaud, T., Delas, M. and Jalier, M. (2004) Key figures of the Ile-de-France Region – 2004. Institute for Urban Planning of the Paris Ile-de-France Region (IAURIF) website.
http://www.iaurif.org/en/doc/studies/key_figures/index.htm.
- Beutler, M., Wiltshire, K.H., Meyer, B., Moldaenke, C., Lüring, C., Meyerhöfer, M., Hansen, U.P. and Dau, H. (2002) A fluorometric method for the differentiation of algal populations *in vivo* and *in situ*. *Photosynthesis Research* 72(1), 39-53.
- Burford, M.A., Johnson, S.A., Cook, A.J., Packer, T.V., Bradley, M.T. and Townsley, E.R. (2007) Correlations between watershed and reservoir characteristics, and algal blooms in subtropical reservoirs. *Water Research* 41(18), 4105-4114.
- Casgrain, P. and Legendre, P. (2004) The R package for multivariate and spatial analysis, version 4.0 d9 – User's manual. Département des sciences biologiques, Université de Montréal, Canada.
<http://www.bio.umontreal.ca/Casgrain/en/>.
- Cheruvilil, K.S., Soranno, P.A., Bremigan, M.T., Wagner, T. and Martin, S.L. (2008) Grouping lakes for water quality assessment and monitoring: the roles of regionalization and spatial scale. *Environmental Management* 41(3), 425-440.
- Chorus, I. and Bartram, J. (1999) Toxic cyanobacteria in water: a guide to public health significance, monitoring and management. Spon /Chapman & Hall, London.
- Chorus, I. (2001) Cyanotoxins. Occurrence, causes, consequences. Springer-Verlag KG, Berlin.
- Colwell, R.K. and Coddington, J.A. (1995) Biodiversity Measurement and Estimation. Hawksworth, D.L. (ed), pp 101-108, Chapman and Hall, London.
- Cook, C.M., Vardaka, E. and Lanaras, T. (2004) Toxic cyanobacterial in Greek freshwaters, 1987-2000: Occurrence, toxicity and impacts in the Mediterranean region. *Acta Hydrochimica et Hydrobiologica Weinheim* 32(2), 107-124.
- Detenbeck, N.E., Cincotta, D., Denver, J.M., Greenlee, S., Olsen, A.R. and Pitchford, A.M. (2005) Watershed-based survey design. *Environmental Monitoring and Assessment* 103(1-3), 59-81.
- Duarte, C. and Canfield, D.E. (1992) Patterns in phytoplankton community structure in Florida lakes. *Limnology and Oceanography* 37(1), 155-161.
- Elliott, J.M. (1971) Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Association Scientific Publication, Westmorland.
- Even, S., Billen, G., Bacq, N., Théry, S., Ruelland, D., Garnier, J., Cugier, P., Poulin, M., Blanc, S., Lamy, F. and Paffoni, C. (2007) New tools for modelling water quality of hydrosystems: An application in the Seine River basin in the frame of the Water Framework Directive. *Science of the Total Environment* 375(1-3), 274-291.
- Frontier, S. (1983) Stratégie d'échantillonnage en écologie. Masson, Paris.
- Galbraith, L.M. and Burns, C.W. (2007) Linking land use, water body type and water quality in southern New Zealand. *Landscape Ecology* 22, 231-241.

- Goslar, T., Ralska-Jasiewiczowa, M., van Geel, B., Lacka, B., Szeroczyńska, K., Chróst, L. and Walanus, A. (1999) Anthropogenic changes in the sediment composition of Lake Gosciadz (central Poland) during the last 330 years. *Journal of Paleolimnology* 22(2), 171-185.
- Gregor, J., Maršálek, B. and Šípková, H. (2007) Detection and estimation of potentially toxic cyanobacteria in raw water at the drinking water treatment plant by *in vivo* fluorescence. *Water Research* 41(1), 228-234.
- Huisman, J., Matthijs, H.C.P. and Visser, P.M. (2005) Harmful cyanobacteria. *Aquatic Ecology Series*, Springer, Dordrecht.
- Hulbert, S.H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54(2), 187-211.
- IFN - Inventaire Forestier National (2007) La forêt française: les résultats issus des campagnes d'inventaire 2005 et 2006. Inventaire Forestier National, Paris.
- Jayatissa, L.P., Silva, E.I.L., McElhiney, J. and Lawton, L.A. (2006) Occurrence of toxigenic cyanobacterial blooms in freshwaters of Sri Lanka. *Systematic and Applied Microbiology* 29(2), 156-164.
- Jones, J.R., Knowlton, M.F., Obrecht, D.V. and Cook, E.A. (2004) Importance of landscape variables and morphology on nutrients in Missouri reservoirs. *Canadian Journal of Fisheries and Aquatic Sciences* 61(8), 1503-1512.
- Knoll, L.B., Vanni, M.J. and Renwick, W.H. (2003) Phytoplankton primary production and photosynthetic parameters in reservoirs along a gradient of watershed land use. *Limnology and Oceanography* 48(2), 608-617.
- Kratz, T.K., Webster, K.E., Bowser, C.J., Magnuson, J.J. and Benson, B.J. (1997) The influence of landscape position on lakes in northern Wisconsin. *Freshwater Biology* 37(1), 209-217.
- Legendre, P., Legendre, L. (1998) *Numerical ecology*. Elsevier, Amsterdam.
- Magurran, A.E. (1995) *Ecological diversity and its measurement*. Princeton University Press, Princeton.
- Mehaffey, M.H., Nash, M.S., Wade, T.G., Ebert, D.W., Jones, K.B. and Rager, A. (2005) Linking land cover and water quality in New York City's water supply watersheds. *Environmental Monitoring and Assessment* 107(1-3), 29-44.
- Marchand, O. (2007) Enquêtes annuelles de recensement de 2004-2006. Insee première 1116, Insee, Paris.
- Morris, C.E., Bardin, M., Berge, O., Frey-Klett, P., Fromin, H., Girardin, H., Guinebretière, M.H., Lebaron, P., Thiéry, J.M. and Troussellier, M. (2002) Microbial biodiversity: approaches to experimental design and hypothesis testing in primary scientific literature from 1975 to 1999. *Microbiology and Molecular Biology Reviews* 66(4), 592-616.
- Olding, D.D., Hellebust, J.A. and Douglas, M.S.V. (2000) Phytoplankton community composition in relation to water quality and water-body morphometry in urban lakes, reservoirs and ponds. *Canadian Journal of Fisheries and Aquatic Sciences* 57(10), 2163-2174.
- Pielou, E.C. (1975) *Ecological diversity*. John Wiley & Sons, New York.
- Reynolds, C.S. and Petersen, A.C. (2000) The distribution of planktonic cyanobacteria in Irish lakes in relation to their trophic states. *Hydrobiologia* 424(1-3), 91-99.
- Rolland, A., Bird, D.F. and Giani, A. (2005) Seasonal changes in composition of the cyanobacterial community and the occurrence

of hepatotoxic blooms in the eastern townships, Québec, Canada. *Journal of Plankton Research* 27(7), 683-694.

Ryding, S.O. and Rast, W. (1989) *The control of eutrophication of lakes and reservoirs*. Unesco and the Parthenon Publishing Group, Paris.

Seaber, P.R., Kapinos, F.P. and Knapp, G.L. (1987) *Hydrologic unit map*. United States Geological Survey Water-Supply Paper 2294, U.S. Geological Survey, Reston.

Southwood, T.R.E. (1978) *Ecological Methods*. Wiley/Halsted, New York.

Talling, J.F. and Driver, D. (1963) *Proceedings on primary production measurement, marine and freshwater*. Doty, M. (ed), pp. 142-146, US Atomic Energy Engineering Commission., Hawaii.

Tong, S.T.Y and Chen, W. (2002) *Modelling the relationship between land use and surface water quality*. *Journal of Environment Management* 66(4), 377-393.

Wasson, J.G., Chandesris, A., Pella, H. and Blanc, L. (2002) *Typology and Classification*

of Lakes and Rivers. Ruoppa, M. and Karttunen, K. (eds), pp 37-41, *Symposium on Typology and Classification of Lakes and Rivers*, Helsinki.

Wiedner, C., Chorus, I. and Fastner, J. (2001) *Cyanotoxins. Occurrence, causes, consequences*. Chorus, I. (ed), pp 49-56, Springer, Berlin.

Willame, R., Jurczak, T., Iffly, J.-F., Kull, T., Meriluoto, J. and Hoffmann, L. (2005) *Distribution of hepatotoxic cyanobacterial blooms in Belgium and Luxembourg*. *Hydrobiologia* 551(1), 99-117.

Willén, T. and Mattson, R. (1997) *Water-blooming and toxin-producing cyanobacterial in Swedish fresh and brackish waters, 1981-1995*. *Hydrobiologia* 353(1-3), 181-192.

Williams, D.A. (1976) *Improved likelihood ratio tests for complete contingency tables*. *Biometrika* 66, 33-37.

Young, W.M., Marston, F.M. and Davis, J.R. (1996) *Nutrient exports and land use in Australian catchments*. *Journal of Environment Management* 47(2), 165-183.