



Patterns and multi-scale drivers of phytoplankton species richness in temperate peri-urban lakes



Arnaud Catherine ^{a,*}, Maloufi Selma ^{a,1}, David Mouillot ^b, Marc Troussellier ^b, Cécile Bernard ^a

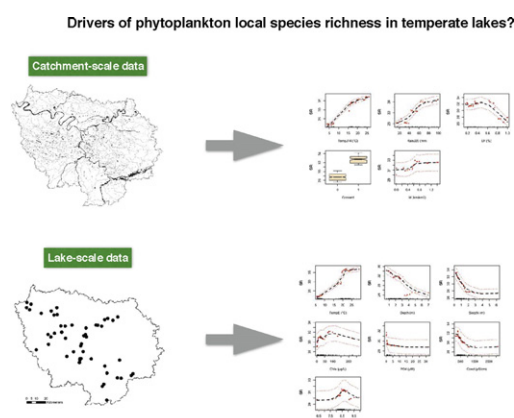
^a UMR7245 MCAM MNHN-CNRS, Muséum National d'Histoire Naturelle, CC 39, 12 rue Buffon, F-75231 Paris, Cedex 05, France

^b UMR 9190 MARBEC UM2-CNRS-IRD-UM1-IFREMER, CC 93, Place Eugène Bataillon, Université de Montpellier 2, F-34095 Montpellier, France

HIGHLIGHTS

- We studied phytoplankton communities in 50 peri-urban lakes.
- We assessed the impact of multi-scale drivers of phytoplankton richness.
- Local- and catchment-scale predictive models performed similarly.
- Seasonal temperature variation and resource availability strongly modulate species richness.
- This approach may be used for the management and conservation of aquatic ecosystems.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 12 November 2015

Received in revised form 7 January 2016

Accepted 24 March 2016

Available online 5 April 2016

Editor: D. Barcelo

Keywords:

Phytoplankton

Species richness

Lake- and catchment-scale

Random forest

Residuals analysis

Productivity–diversity relationship

ABSTRACT

Local species richness (SR) is a key characteristic affecting ecosystem functioning. Yet, the mechanisms regulating phytoplankton diversity in freshwater ecosystems are not fully understood, especially in peri-urban environments where anthropogenic pressures strongly impact the quality of aquatic ecosystems. To address this issue, we sampled the phytoplankton communities of 50 lakes in the Paris area (France) characterized by a large gradient of physico-chemical and catchment-scale characteristics. We used large phytoplankton datasets to describe phytoplankton diversity patterns and applied a machine-learning algorithm to test the degree to which species richness patterns are potentially controlled by environmental factors. Selected environmental factors were studied at two scales: the lake-scale (e.g. nutrients concentrations, water temperature, lake depth) and the catchment-scale (e.g. catchment, landscape and climate variables). Then, we used a variance partitioning approach to evaluate the interaction between lake-scale and catchment-scale variables in explaining local species richness. Finally, we analysed the residuals of predictive models to identify potential vectors of improvement of phytoplankton species richness predictive models.

Lake-scale and catchment-scale drivers provided similar predictive accuracy of local species richness ($R^2 = 0.458$ and 0.424 , respectively). Both models suggested that seasonal temperature variations and nutrient supply strongly modulate local species richness. Integrating lake- and catchment-scale predictors in a single predictive

* Corresponding author.

E-mail addresses: arnocat@mnhn.fr (A. Catherine), maloufi@mnhn.fr (M. Selma), david.mouillot@univ-montp2.fr (D. Mouillot), troussel@univ-montp2.fr (M. Troussellier), cbernard@mnhn.fr (C. Bernard).

¹ These authors contributed equally to this work.

model did not provide increased predictive accuracy; therefore suggesting that the catchment-scale model probably explains observed species richness variations through the impact of catchment-scale variables on in-lake water quality characteristics.

Models based on catchment characteristics, which include simple and easy to obtain variables, provide a meaningful way of predicting phytoplankton species richness in temperate lakes. This approach may prove useful and cost-effective for the management and conservation of aquatic ecosystems.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Spatial biogeographic patterns are increasingly studied in order to provide a better understanding of the ecology of living organisms (e.g. Marquet et al., 2004; Azeria et al., 2009) and to identify the processes involved in the maintenance or decline of biodiversity (Bellwood and Hughes, 2001). Evidence on the occurrence of biogeographic patterns in aquatic microorganisms has only recently been provided (Foissner, 2006; Martiny et al., 2006; Green and Bohannan, 2006) and the deterministic or random nature of the underlying processes of microbial biodiversity is still under close scrutiny. However, understanding how the environment acts on biodiversity patterns in microorganisms is critical for mitigating the impact of environmental changes and ensuring the continuity of ecosystems services (Millennium Ecosystem Assessment, 2005). This is particularly true for phytoplankton communities, where shifts in communities' composition (e.g. leading to species poor communities dominated by harmful cyanobacteria) may have profound effects on aquatic ecosystem functioning and on the quality of aquatic resources.

Various processes have been suggested to impact microorganisms' richness patterns across scales, including latitudinal temperature gradients at large spatial scale (Fuhrman et al., 2008), species-area relationships (e.g. Horner-Devine et al., 2004; Smith et al., 2005; Bell et al., 2005), temporal (i.e. stability; Ptacnik et al., 2008b) and spatial (e.g. water column stratification; Streibel et al., 2010) ecosystem heterogeneity and the size of the regional pool of potential colonizers (Ptacnik et al., 2010). Among the various processes linking diversity to ecosystem functioning, the relationship between diversity and productivity (Currie, 1991) has been actively debated in the last decades (Strong, 2010). In essence, productivity corresponds to the ratio of production over biomass and characterizes the efficiency of a biological compartment to use surrounding resources. However, most empirical studies use standing biomass as a surrogate measure of productivity (e.g. Groner and Novoplansky, 2003; Filstrup et al., 2014; Vallina et al., 2014). Current knowledge suggests that the shape of productivity–diversity relationships in both terrestrial and aquatic ecosystems is either positive or hump-shaped (e.g. Dodson et al., 2000; Mittelbach et al., 2001; Chase and Leibold, 2002; Gillman and Wright, 2006; Smith, 2007; Filstrup et al., 2014; Vallina et al., 2014). A number of hypotheses have arisen in the literature to explain how productivity might drive diversity (Palmer, 1994), including the intermediate disturbance hypothesis (Connell, 1978), the species-energy theory (Wright, 1983), the resources-supply ratios hypothesis (Tilman, 1985) or the keystone-predation hypothesis (Leibold, 1996). Alternatively, it was also suggested that diversity might drive productivity (Loreau et al., 2002; Duffy, 2009). These two views on the relationship between diversity and productivity are currently seen as complementary rather than mutually exclusive (Cardinale et al., 2009a, 2009b).

In most aquatic ecosystems, productivity (and standing phytoplankton biomass) is at least partly controlled by resource availability (Vallina et al., 2014). The occurrence of a productivity–diversity relationship in phytoplankton advocates a deterministic control of resource availability on local species richness (SR) and is supported by recent findings (Cardinale et al., 2009a, 2009b). However, to date, most ecological studies on phytoplankton diversity patterns have traditionally focused on understanding among sites variations in SR using explanatory variables quantified at relatively fine scales

(e.g. in-lake nutrient concentration). While these studies have provided valuable approaches to test functional hypotheses regarding the drivers of phytoplankton species richness, (i) the hydrogeomorphic (e.g. hydrological connectivity) and anthropogenic features (e.g. land use) occur at multiple scales (Levin, 1992; Turner et al., 2001) and (ii) their interactions with meteorological factors have been seldom studied.

In this study, we examined species richness patterns of phytoplankton communities across 50 freshwater water bodies located in the Paris area (within a 200 km radius), a region characterized by strong gradients in anthropogenic pressure and by different degree of hydrological connectivity of water bodies (Catherine et al., 2008, 2010). We assessed the role of lake-scale and catchment-scale variables in explaining variations of species richness in space and time. Finally, we analysed the residuals of predictive models in order to identify vectors of improvement of predictive models of phytoplankton richness in temperate lakes.

2. Materials and methods

2.1. Study area and sample collection

The Paris area (Fig. 1) is situated in north-central France and constitutes the most densely populated French region and is home to 11.9 million people (19% of the national population; IAURIF, 2014). The region includes many large industrial towns and residential suburbs. However, agricultural and forested areas still cover over 50% and 23% of the region, respectively (IAURIF, 2014), leading to contrasted local environmental contexts. The Paris area includes around 248 lakes and ponds covering over 5 ha (Catherine et al., 2008), most of which are old sand and gravel quarries that were worked between the 1940s and 1980s, or that resulted from peat extraction in the mid-19th century. The remaining water bodies are reservoirs built in the 17th and 18th centuries to provide a reliable water supply for Versailles Castle and the city of Paris. Nowadays, these water bodies support various human activities including water sports and other recreational activities, and fishing.

Data were collected from 50 lakes and reservoirs, which were selected according to a stratified sampling strategy previously described in Catherine et al. (2008). Most selected water bodies are shallow (average depth of water bodies in the area <2.5 m) and represent a wide range of environmental conditions (e.g. size of watershed, land use, hydrology and trophic status) (Catherine et al., 2010).

Sampling was conducted over two summers (50 lakes in August 2006 and 48 lakes in August 2011, hereafter termed Aug06 and Aug11) and one winter (50 lakes in February 2007, hereafter termed Feb07). Each sampling campaign was performed over a short timeframe (approximately two weeks) to obtain a 'snapshot' view of the phytoplankton communities in the selected lakes. Three stations were sampled in each lake to integrate spatial heterogeneity. For each sampling station, a 5 L sample was obtained at three depths (i.e. 0.5–1 m below the surface and at two depths equally spaced according to the lake's depth at the sampling stations) using a Niskin water sampler (KC Denmark, Silkeborg, Denmark). All nine samples per lake were pooled and used for phytoplankton community composition analysis and water chemistry measurements.

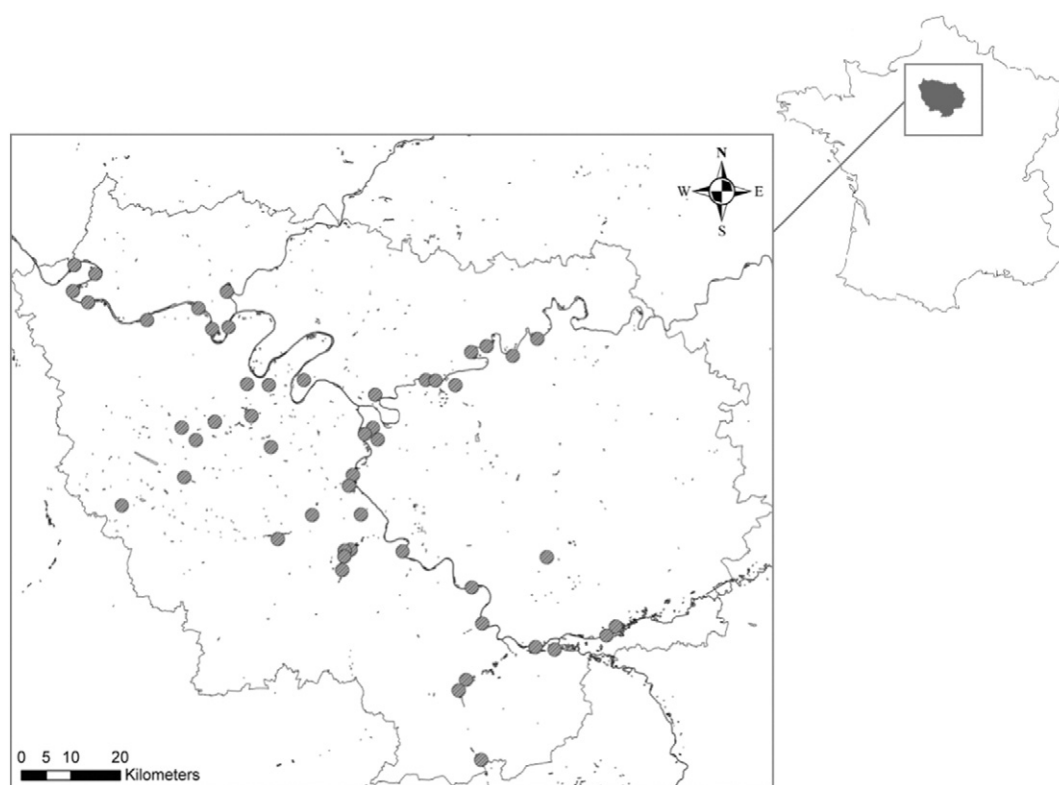


Fig. 1. Map of the Paris area (Ile-de-France region) with the location of the studied lakes ($n = 50$).

2.2. Phytoplankton species data

Following phytoplankton identification, counts data were obtained using the Utermöhl method under an Eclipse TS100 inverted microscope at $600\times$ magnification (Nikon Instruments Inc., Melville, NY, USA) as described in EN 15204:2006 (2006). Samples were acclimatised to room temperature before settling to facilitate a homogeneous and random distribution of algal cells at the bottom of the settling chamber. Depending on the density of phytoplankton cells, 5–18 mL of each sample was transferred into a 20 mL Utermöhl settling chamber (Olympus Optical Co., Tokyo, Japan). For samples with high biomass (chlorophyll $a > 5 \mu\text{g L}^{-1}$), more than 30 random fields were counted, corresponding to more than 500 algal units (algal cell, colony or filament/trichome) per replicate. For samples with low biomass (chlorophyll $a < 5 \mu\text{g L}^{-1}$), up to 200 random fields were counted, corresponding to over 250 algal units per replicate. Colonial phytoplankton species (mainly *Microcystis* spp.) were spread apart using a VibraCell VCX130 (Sonics, Newtown, CT USA)

ultrasonic probe (two 5 s pulses at 50%) and counted separately. Count data were then transformed into presence–absence and presences were summed to determine species richness per lake (SR). This method was shown to provide accurate estimates of local species richness (Maloufi et al., in press).

2.3. Lake-scale (LS) and catchment-scale (CS) predictor variables

Environmental predictor variables were selected based on their known direct or indirect impact on ecosystem productivity (Ptacnik et al., 2008a; Catherine et al., 2010; Stomp et al., 2011; Soinen and Luoto, 2012; Özkan et al., 2013). Two sets of variables, corresponding to either the lake-scale (LS) or the catchment-scale (CS), were selected to evaluate their ability to predict local SR (Table 1).

All samples used for nutrient analysis were collected in acid-washed polyethylene containers. Samples were immediately filtered using a cellulose acetate $0.22\text{-}\mu\text{m}$ syringe filter (Nalgene, Rochester, NY) for dissolved nutrient analysis. Ammonium (NH_4^+) and orthophosphate

Table 1

Selected explanatory of variables of the lake-scale (LS model) and catchment-scale (CS model) predictive models of phytoplankton species richness.

Lake-scale			Catchment-scale		
Variable	Description	Unit	Variable	Description	Unit
TP	Total phosphorus	μM	WB _a	Water body surface area	ha
TN	Total nitrogen	μM	Catch _a	Catchment surface area	km^2
Chl _a	Chlorophyll <i>a</i>	$\mu\text{g/L}$	Connect	Connectivity to the hydrological network	-
NO_3^-	Nitrate	μM	I_d	Density of water ways in the catchment	km/km^2
NH_4^+	Ammonium	μM	Alt	Altitude of water body	m
PO_4^{3-}	Orthophosphate	μM	LA	Catchment surface covered by agricultural surfaces	%
Secchi	Water transparency	m	LI	Catchment surface covered by impervious surfaces	%
Cond	Water conductivity	$\mu\text{S/cm}$	LF	Catchment surface covered by forests	%
Temp _E	Water temperature	$^\circ\text{C}$	Rain ₂₀	Cumulative rainfall over 20 days prior to sampling	mm
pH	-	-	Rain ₀	Rainfall on the day of sampling	mm
Depth	Mean depth of water body	m	Temp ₁₄	Average air temperature over 14 days prior to sampling	$^\circ\text{C}$

(PO_4^{3-}) analyses were carried out on the day of sampling using colorimetric methods previously described in Eaton et al. (2005). Nitrate (NO_3^-) was measured using a DX600 ion chromatograph (Dionex Corp., Westmont, IL) equipped with an AS14 Ion Pack analytical column (Dionex Corp., Westmont, IL). TP was measured by colorimetry following ammonium persulfate digestion (Eaton et al., 2005). TN was measured by UV-spectroscopy following potassium persulfate digestion (Eaton et al., 2005). Water temperature, pH and conductivity were measured using a multiparameter Sea-Bird SBE 19 Seacat Profiler (Sea-Bird Electronics Inc., WA). Chlorophyll *a* concentration (Chl*a*) was measured fluorometrically using a FluoroProbe II device (bbe-Moldaenke, Kiel, Germany) as described in Catherine et al. (2012).

Catchment delineation was performed on the basis of digital altitudinal data using a sampling rate of one point every 50 m and obtained from the BD Alti database (IGN, France). Altitudinal data were handled using Vertical Mapper 3.1 (TETRAD Computer Applications Inc., Bellingham, WA, U.S.A.), an add-on to MapInfo 8.0 Professional (MapInfo Inc., Troy, NY, U.S.A.). Delineation was corrected using a 1:25,000 detailed geographic map (IGN, France), to allow for artificial linear structures (e.g. roads, ditches, culverts), which act as artificial boundaries. Water bodies' surface area (in km^2), their connectivity to the hydrological network (binary variable) and the drainage index (i.e. the density of waterways within their catchment in km/km^2) were extracted from the BD Carthage hydrological database (IGN, France). Land use data were extracted from the MOS 2009 database (IAURIF, France). All meteorological data were acquired from the French national meteorological agency (MeteoFrance). For each lake, the weather station nearest to the water body was selected (usually less than 10 km away).

2.4. Statistical analysis

The relationship between each set of predictor variables (i.e. lake- and catchment-scale; hereafter called LS and CS model, respectively) and local species richness (SR) was modeled using the Random Forest (RF) algorithm (Breiman, 2001) implemented in the 'randomForest' package (Liaw and Wiener, 2002). The data from the three sampling campaigns were combined for model fitting and for model performance evaluation. RF is an ensemble learning method that builds a large collection of classification or regression trees. As our response variable is quantitative, we used RF in regression mode. Trees are built using a bootstrap sample of the observations and a random set of potential predictors (i.e. explanatory variable) to determine the best split at each node. Trees are grown to maximum size without pruning, and the aggregation of trees is performed by averaging. Variables importance in RF models was evaluated using function 'varimpAUC' (package 'party'; Janitza et al., 2013), an AUC-based permutation variable importance measure that provides better performance than alternative methods when the distribution of predictors' values is strongly skewed (Janitza et al., 2013). The RF algorithm is considered as being generally insensitive to collinearity among predictor variables. However, including multicollinear predictors may impact variable importance estimates and, in some cases, reduce the predictive accuracy of RF models (Murphy et al., 2010). The selected sets of predictors were thus tested for multicollinearity using QR-matrix decomposition (function 'multi.collinear' from package 'rftUtilities'; Evans and Murphy, 2015). The method did not identify any significant multivariate redundancy in either of the two sets of predictor variables. We then used a modified version of the variable selection method described in Murphy et al. (2010), where classical variable importance measures (based on out-of-bag error rate) was replaced by AUC-based variable importance measures, in order to select for the most parsimonious model. The predictive accuracy of the final RF models was evaluated through Monte Carlo cross-validation (Mercier et al., 2011), where randomly selected subsets of the data (80% of observations) were used for fitting the models and

the remaining data (20% of observations) for evaluating the predictive power of the models (through the R^2 and MSE scores). This procedure was repeated 100 times, and the predictions obtained were averaged for each model. Finally, we evaluated the marginal effects of predictor variables in the final models using the function 'partial.plot' (package 'randomForestSRC'; Ishwaran and Kogalur, 2014).

In order to evaluate if both model explained independently (or not) local SR a variance partitioning approach was used. In short, after pulling the predictors included in the LS and CS model, an increase in predictive power of the combined model would therefore mean that, to some extent, both models explain independently SR values. At the opposite, a lack of increase in predictive power, would suggest that merging the two models led to redundant information and that the catchment-scale model probably explains observed SR variations through the impact of catchment-scale variables on in-lake water quality characteristics. This procedure is adequate for random forests as the rate of convergence of this algorithm is not influenced by how many noise variables are present in the model (Biau, 2012).

Finally, we analysed the occurrence of spatial autocorrelation in the residuals of the LS and CS predictive models. Residual spatial autocorrelation (RSA) violates the assumption of independent and identically distributed errors and can inflate type I errors (Legendre, 1993; Kühn, 2007). RSA might originate from random model error but also from unexplained natural processes. The analysis of the residuals of predictive models may thus provide meaningful insights for improving predictive models (e.g. identification of spatially structured environmental factor unaccounted for on the initial model). The occurrence of RSA was tested using distance-based Moran's eigen-vector maps (dbMEM also called PCNM; Borcard and Legendre, 2002). dbMEM eigenfunctions were obtained by principal coordinate analysis (PCoA) of a truncated geographic distance matrix among the sampling sites (Borcard and Legendre, 2002; Borcard et al., 2004). Only positive eigenfunctions were selected as explanatory variables to analyse the spatial variation of the phytoplankton community composition data (Legendre and Legendre, 2012). When the full model (i.e. including all positive dbMEM eigenvectors) was significant, a forward selection (with permutation tests, at the 5% significance level, of the increase in R^2 at each step) was applied to the eigenfunctions tables in order to avoid overfitting (Borcard et al., 2011). The proportion of variance in community composition data explained by eigenvectors was assessed using canonical redundancy analysis (RDA; Rao, 1964).

All analyses were implemented in the R environment 3.1.1. (R Core Team, 2014).

3. Results

3.1. Local and regional species richness patterns

Higher regional species richness (γ -diversity) values were found in summer ($\gamma_{\text{Aug06}} = 340$; $\gamma_{\text{Aug11}} = 383$) compared to winter ($\gamma_{\text{Feb07}} = 244$) (Table 2). Higher mean local SR values in summer than in winter were also observed (mean SR = 38, 35 and 19 for the Aug06, Aug11 and Feb07 sampling campaign, respectively). In term of phytoplankton community composition, the region is characterized by high temporal variability, with only 76 species (out of 676) shared among the three sampling campaigns and about half of summer species were specific

Table 2
Local and regional of phytoplankton species richness.

	August 2006	February 2007	August 2011
Regional species richness	340	244	383
Local species richness			
Min	9	4	5
Max	73	35	70
Mean	38	19	35
CV (%)	42	37	44

of a given sampling campaign (198 and 185 species for Aug06 and Aug11, respectively; data not shown but see Maloufi et al., in press for details). Local SR estimates displayed significant inter-annual (Aug06 vs. Aug11: $r = 0.574$, $p\text{-value} < 0.001$) and inter-seasonal (mean summer vs. winter: $r = 0.447$, $p\text{-value} < 0.001$) correlations, which suggests that the water bodies displaying high species richness tend to remain species rich (and respectively), and this, despite the inter-seasonal variations in mean species richness.

3.2. Predictive accuracy of the lake-scale and catchment-scale models

Available catchment and limnological variables (Table 1 and see Appendix for summary statistics) were first ordered according to their importance in explaining SR values by using the AUC-based variable importance measure (Table 3). Following the variable selection procedure, 7 variables were selected for the lake-scale model (LS model) and 5 variables for the catchment-scale model (CS model). The variables selected in the LS model (Table 3) included water temperature ($Temp_E$), mean depth of water bodies (depth), water transparency (Secchi), chlorophyll *a* concentrations (Chla) and orthophosphate concentrations (PO_4^{3-}), conductivity (Cond) and pH. For the most parsimonious CS model, the 5 variables selected included air temperature (average of over 14 days prior to sampling: $Temp_{J14}$), cumulative rainfall over 20 days prior to sampling ($Rain_{J20}$), the connectivity of water bodies to the hydrographic network (Connect), the percentage of forested cover within catchment (LF), and the density of streams within the catchments (Id).

Both models displayed a similar predictive accuracy (Table 3). For the calibration dataset, both models were able to explain a large proportion of the variance in SR ($R^2_{FIT} = 0.923$ and 0.869 for the LS and CS model, respectively; Table 3). Under cross-validation conditions, the both models provided a good prediction of SR values ($R^2_{CV} = 0.458 \pm 0.106$ and 0.424 ± 0.114 ; $RMSE_{CV} = 12.3 \pm 1.4$ and 12.7 ± 1.7 for the LS and CS model, respectively; Table 3). Pulling all selected predictors from the two models only marginally increased the predictive power of the final model ($R^2_{CV} = 0.472 \pm 0.118$; $RMSE_{CV} = 12.4 \pm 1.6$; Table 3).

Table 3

AUC-based variable importance measure and Random Forest model accuracy of the lake-scale (LS model) and the catchment-scale (CS model) predictive models of phytoplankton species richness.

LS model		CS model	
VarImpAUC		VarImpAUC	
$Temp_E^*$	105.01	$Temp_{J14}^*$	91.36
Depth*	39.27	$Rain_{J20}^*$	26.11
Secchi*	15.24	Connect *	19.91
Chla*	3.31	LF*	8.972
PO_4^{3-} *	2.13	I*	1.40
Cond*	1.62	Catcha	1.36
pH*	1.53	Alt	0.18
NO_3^-	0.75	LA	0.04
TP	0.33	$Rain_{J0}$	-0.19
NH4	-0.12	LI	-0.58
TN	-0.35	WBa	-0.77
R^2_{FIT}	0.923	R^2_{FIT}	0.869
$RMSE_{FIT}$	5.17	$RMSE_{FIT}$	6.20
R^2_{CV}	0.458 ± 0.106	R^2_{CV}	0.424 ± 0.114
Mean $RMSE_{CV}$	12.3 ± 1.4	$RMSE_{CV}$	12.7 ± 1.7
LS + CS model			
R^2_{CV}	0.472 ± 0.118		
Mean $RMSE_{CV}$	12.4 ± 1.6		

Predictors marked with an asterisk correspond to selected variables in the final model. R^2_{FIT} and $RMSE_{FIT}$ correspond to the fit of the model to observed data. R^2_{CV} and $RMSE_{CV}$ are the mean R^2 and mean RMSE of the k-fold cross-validation procedure along with their standard deviation.

3.3. Marginal effects of environmental variables on local species richness

The relationship between the selected explanatory variables and SR values were evaluated using partial dependence plots (Figs. 2 & 3). Temperature had the strongest explanatory power in both models (water and air temperature for the LS and CS model, respectively). In the Limnological model, water temperature ($Temp_E$) displayed an s-shaped relationship with SR values, with increasing SR values up to 19°C (Fig. 2). A very similar relationship with air temperature ($Temp_{J14}$) was found for the CS model (Fig. 3).

Regarding the LS model (Fig. 2), the mean depth of water bodies (Depth) displayed a negative relationship with SR. Increasing water transparency (Secchi), conductivity (Cond) and orthophosphate availability (PO_4^{3-}) were all associated with lower SR, all of these variables displaying decreasing saturating relationship with SR values. The relationship between chlorophyll *a* concentrations (Chla) and SR was hump-shaped with SR decreasing for Chla values over c.a. $100\text{ }\mu\text{g/L}$. A similar hump-shaped relationship was found between SR and pH.

Regarding the CS model, the cumulative rainfall over 20 days prior to sampling ($Rain_{J20}$) and the density of waterways in the catchment (Id) both displayed a positive s-shaped relationship with SR. In addition, water bodies connected to the hydrographic network were found to display higher species richness (Fig. 3). Finally, we found catchment dominated by forested cover (>60% of catchment area) to be associated with lower SR.

3.4. Analysis of models' residuals

The residuals of the LS and CS models (Fig. 4) did not display any significant spatial structure (Table 4). Based on the RDA analysis of dbMEM eigenvectors, none of the model was significant ($p\text{-value} < 0.05$) whatever the model (i.e. LS or CS model) and the sampling campaign considered. However, we found that both models significantly overestimated phytoplankton species richness in winter (Fig. 4; Table 5). In contrast, there was not significant bias in model predictions for the summer data, except for the LS model applied to the summer 2006 data where it slightly underestimated true species richness (Table 5). This pattern did not originate from an unbalance number of observations between summer and winter data. Both models were run using the winter 2007 data along with the data from either the summer 2006 or summer 2011 sampling and model residuals displayed the same patterns (data not shown).

4. Discussion

4.1. Phytoplankton diversity in the Paris area

Local species richness (SR) exhibited major among-lakes variations whatever the sampling campaign considered (Table 2). In addition, we found strong seasonal variations in both regional and mean local species richness (Table 2). The observed regional phytoplankton richness in the Paris area ($\gamma = 676$ species, $n = 50$ lakes) is at the upper range of reported values (Nabout et al., 2007; de Nogueira et al., 2010; Angeler and Drakare, 2013; de Barros et al., 2013). In a previous study (Maloufi et al., in press), we showed that the high γ -diversity of the Paris area originated from hyper β -diversity of phytoplankton communities (>95% of the γ -diversity) in response to environmental heterogeneity across the region.

4.2. Seasonal temperature variations strongly modulate local species richness

In both the lake-scale (LS model) and the catchment-scale (CS model) models, temperature was the parameter most strongly associated with species richness (SR) variations (water temperature ($Temp_E$) and mean air temperature over 14 days prior to sampling ($Temp_{J14}$) for the LS and

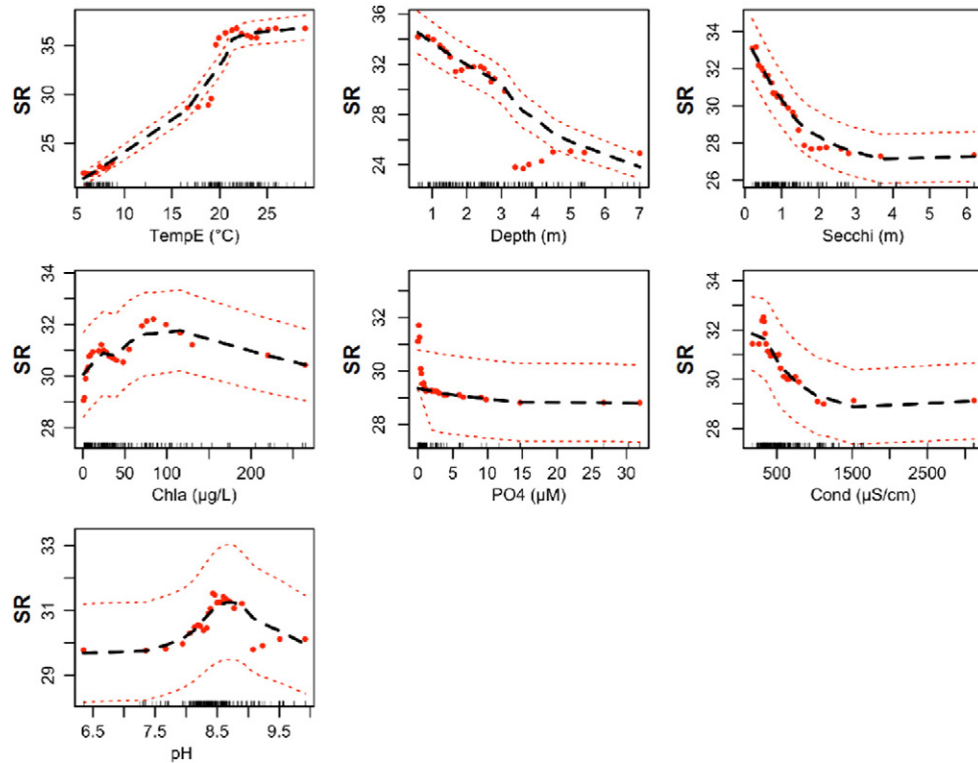


Fig. 2. Partial plot representing the marginal effect of single variables included in the lake-scale random forest model (LS model) on overall local species richness (SR) while averaging out the effect of all the other variables. In a partial plot of marginal effects, only the range of values and (and not the absolute values) can be compared between plots of different variables (TempE: water temperature (°C); Depth: mean depth of water body (m); Secchi: water transparency (m); Chla: chlorophyll *a* concentration (µg/L); PO4: orthophosphate concentration (µM); Cond: water conductivity (µS/cm)). The black dashed line corresponds to a lowess smoothed line representing the partial dependence between each explanatory variable and the response. The dashed red lines indicate a smoothed error bar of \pm two standard errors. The red dots ($n = 25$) correspond to the partial values used to fit the lowess function.

CS models, respectively. Both variables displayed very similar relationships to SR, with species richness increasing roughly linearly at temperature up to 19 °C while higher temperatures did not further increase SR (Figs. 2 & 3). Such clear seasonal pattern on species richness has seldom been described in the literature (but see Eloranta, 1993) and never among an entire landscape of lakes. Available large-scale data relating local phytoplankton species richness in lakes to temperature is limited to the observation of large-scale latitudinal gradient, which is frequently taken as a proxy of mean annual temperature (Stomp et al., 2011).

A number of hypotheses have been proposed to explain why warmer conditions may lead to higher diversity. Among these hypotheses, the ‘physiological tolerance hypothesis’ (Currie et al., 2004) states that species can only persist under certain temperature conditions and is backed by the occurrence of very distinct temperature tolerance among phytoplankton species (e.g. Butterwick et al., 2005). Temperature may also influence diversity by allowing a greater range of energetic lifestyles at warmer temperatures (i.e. the ‘metabolic niche hypothesis’; Clarke and Gaston, 2006). The ‘metabolic niche hypothesis’ postulates that many metabolic lifestyles are viable at optimal temperatures, but that only species having low energy requirements at extreme temperatures can cope with the extra energy expended for stress tolerance.

The water column temperature of the shallow lakes sampled in this study sharply decreases in winter (7.2 ± 1.2 °C for Feb07) compared to summer temperatures (23.5 ± 1.8 °C and 19.8 ± 1.6 °C for Aug06 and Aug11, respectively; data not shown). At such low temperature, fewer species are able to reach sufficiently high growth rates to compensate for cell losses originating from predation, cell death or cell sedimentation (Butterwick et al., 2005; Reynolds, 2006; Bissinger et al., 2008), which may explained the lower observed SR values in winter.

4.3. Productivity–diversity relationships and the role of resources limitation

Based on the lake-scale model (LS model), we found a hump-shaped relationship between phytoplankton species richness (SR) and Chla concentrations. While SR increased as a function of Chla concentrations up to 100 µg/L, more productive lakes then tend to display reduced SR values (Fig. 2). Previous studies have shown that phytoplankton diversity displays a positive or hump-shaped relationship with phytoplankton productivity (using Chla concentrations as a surrogate measure), the shape of the relationship depending on the productivity gradient considered (Dodson et al., 2000; Mittelbach et al., 2001; Irigoien et al., 2004; Stomp et al., 2011). Our data confirmed previous observation by Stomp et al. (2011), where, based on a dataset of 540 lakes and reservoirs distributed across the continental United States, a decrease in species richness was observed for Chla concentrations greater than 100 µg/L.

Overall, it is important to note that the marginal effect of productivity itself appears to be rather limited (Fig. 2). Indeed, the LS model also showed that other processes are likely linked to observed species richness variation. Conductivity, despite having a moderate impact on local phytoplankton richness (Fig. 2), appears to be negatively correlated to species richness, which may be related to varying species tolerance to conductivity even in the range of 0–1000 µS/cm (Tibby and Reid, 2004). Most importantly, the model also supported the hypothesis that competition among phytoplankton species for limiting resources strongly impacts phytoplankton SR. Indeed, all else being equal low orthophosphate and low light (i.e. low transparency) availability were shown to be associated to higher species richness. Phosphorous limitation in the lakes sampled in this study is highly likely considering that 95% of observations had TN:TP > 16 (data not shown). In addition, the fact that these lakes are highly productive suggests that

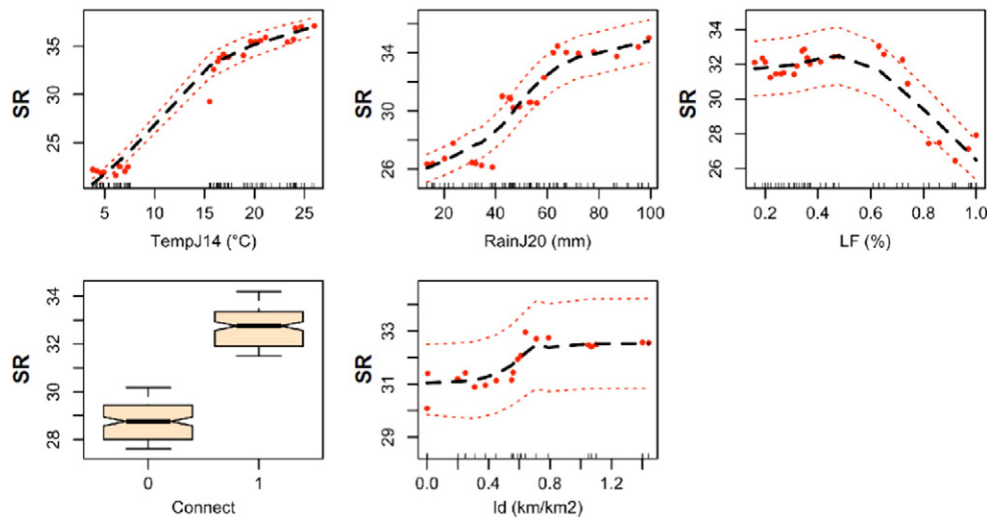


Fig. 3. Partial plot representing the marginal effect of single variables included in the catchment-scale random forest model (CS model) on overall local species richness (SR) while averaging out the effect of all the other variables. In a partial plot of marginal effects, only the range of values and (and not the absolute values) can be compared between plots of different variables (TempJ14: average air temperature (°C) over 14 days prior to sampling; RainJ20: cumulative rainfall (mm) over 20 days prior to sampling; LF: catchment surface (%) covered by forested areas; Connect: connectivity to the hydrographical network (0 and 1 correspond to isolated ($n = 24$) and connected ($n = 26$) water bodies, respectively); Id: density of water ways (km/km²) in the catchment). The black dashed line corresponds to a lowess smoothed line representing the partial dependence between each explanatory variable and the response. The dashed red lines indicate a smoothed error bar of \pm two standard errors. The red dots ($n = 25$) correspond to the partial values used to fit the lowess function.

light availability may often strongly impact phytoplankton productivity. The negative marginal effect of the mean depth of lakes on species richness may also result from interactions between competing primary producers (i.e. between phytoplankton & macrophytes). Indeed, Fig. 2. suggests that phytoplankton richness decreases sharply in lakes having a mean depth greater than 3 m, which corresponds to the maximum colonization of most macrophytes in low transparency lakes (Middelboe and Markager, 1997). Another explanation could

arise from the fact that macrophyte occurrences might induce spatial heterogeneity in physico-chemical conditions, allowing the coexistence of a greater diversity of phytoplankton species.

Finally, pH appears to have a moderate impact on local phytoplankton richness (Fig. 2). However, the processes by which pH might impact diversity are still unclear. In shallow productive lakes, pH is controlled by the relative rates of respiration and photosynthesis (Wetzel, 2001) and high pH is thus often related to high primary production. The

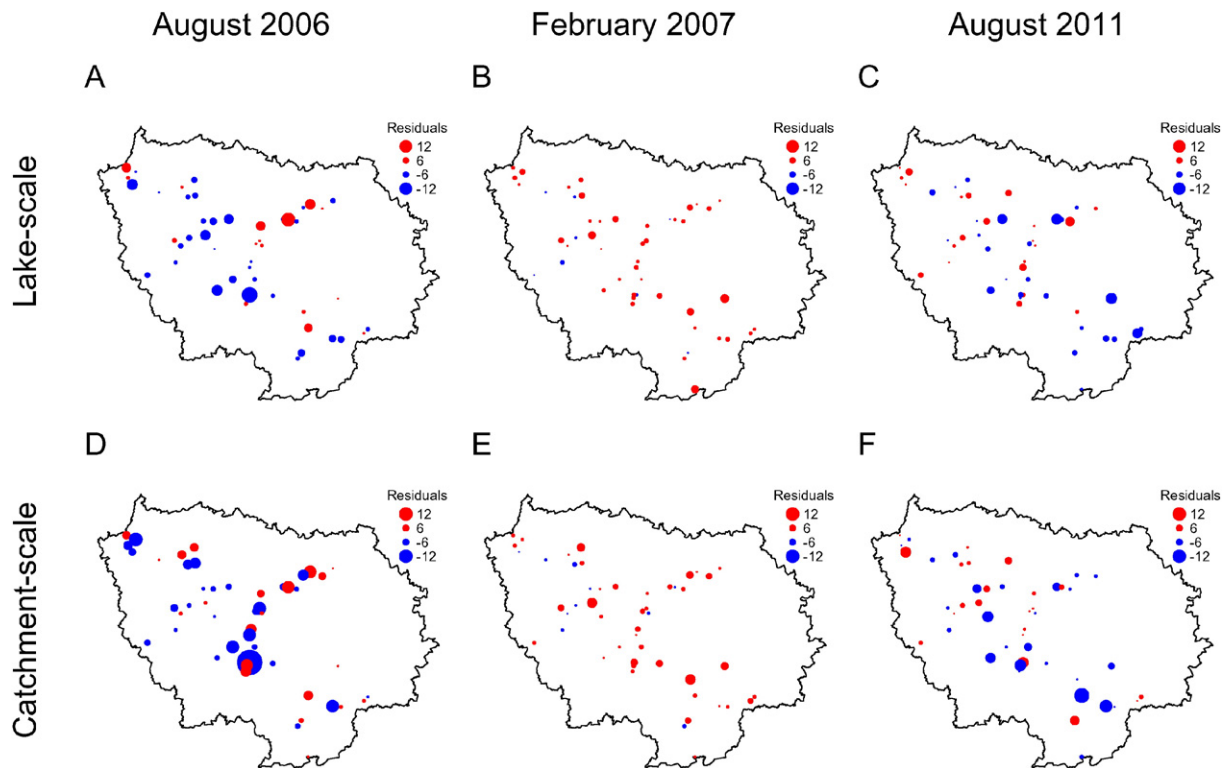


Fig. 4. Maps showing the spatial distribution of residuals of the lake-scale (LS model; A–C) and the catchment-scale (CS model; D–F) predictive models of local species richness (A, D: August 2006; B, E: February 2007; C, F: August 2011). Blue and red circles correspond to lakes where the model underestimated or overestimated, respectively, the true species richness. The size of the circles is proportional to the error of the model.

Table 4

Anova on dbMEM-based RDA analyses of the residuals of the lake-scale (LS model) and the catchment-scale (CS model) predictive models of phytoplankton species richness.

Sampling			Df	Var	F	p	R ² adj
LS model	August 2006	Model	17	8.47	0.970	0.493	0.00
		Residuals	32	16.44			
	February 2007	Model	17	2.46	1.92	0.055	0.24
		Residuals	32	2.40			
	August 2011	Model	17	5.43	0.9356	0.545	0.02
		Residuals	30	10.23			
CS model	August 2006	Model	17	9.76	0.400	0.966	0.00
		Residuals	32	46.01			
	February 2007	Model	17	4.04	1.511	0.169	0.15
		Residuals	32	5.03			
	August 2011	Model	17	7.52	0.653	0.820	0.00
		Residuals	30	20.34			

No forward selection of dbMEM eigenvectors was performed as none of the full models were significant.

observed relationship may originate from some species having a lower tolerance towards extreme pH values or to low CO₂ availability.

Recently Vallina et al. (2014) hypothesized that, the hump-shaped productivity–diversity relationship for phytoplankton might originate from (i) competitive exclusion mechanisms among phytoplankton species originating from resources limitation and (ii) predation mediated species coexistence mechanisms occurring in the middle range of the productivity spectrum. In the lower range of the productivity spectrum, low resource availability should lead to competitive exclusion and decreased species richness. Under these conditions, phytoplankton standing biomass is expected to be too low to sustain significant zooplankton populations. This tends to favour species coexistence as dominant prey species have a greater probability of being grazed. In the higher range of the productivity spectrum, fast growing or inedible bloom-forming opportunist species prevent predation mediated species coexistence, which leads to competitive exclusion and reduced SR. In this study, the marginal effect of low resource availability was associated to increased species richness (Fig. 2), because most sampled lakes are sufficiently productive to allow significant zooplankton population to grow (Stomp et al., 2011) leading to predation mediated coexistence. Accordingly, our data strongly support that, in highly eutrophic lakes, resources limitation allows more diverse phytoplankton communities and provide further evidence that the occurrence of multiple resources limitation (e.g. light, nutrients) in lakes favours higher phytoplankton species richness (e.g. Interlandi and Kilham, 2001).

However, we also found that both the LS and CS predictive models tend to overestimate observed SR in winter (Fig. 4 and Table 5), suggesting that some unexplained process drives richness down in winter. Considering previous hypotheses, this observation could originate from low predation pressure by zooplankton in winter which strengthen competitive exclusion when temperatures are too low to sustain significant zooplankton populations. Further model development should seek at integrating more explicitly biotic processes related to competitive exclusion and predation mediated coexistence in order to increase the performance of predictive SR models.

Table 5

One-sample t-test on the residuals of the lake-scale (LS model) and the catchment-scale (CS model) predictive models of phytoplankton species richness.

		Residuals		One-sample t-test		
		Mean	sd	t	df	p-Value
LS model	August 2006	−1.62	5.06	−2.270	49	0.028
	February 2007	2.34	2.30	7.188	49	< 0.0001
	August 2011	−0.75	4.14	−1.252	47	0.217
CS model	August 2006	−1.32	7.37	−1.271	49	0.210
	February 2007	2.31	2.92	5.589	49	< 0.0001
	August 2011	−1.02	5.16	−1.374	47	0.176

In all cases, the residuals were normally distributed (Anderson-Darling test).

4.4. Catchments characteristics as integrators of biodiversity processes

In addition to the modulating effect of seasonal temperature variations on local phytoplankton richness, our results support the fact that lakes trophic status and competition for limiting resources strongly modulate local species richness. Nutrient supply has a direct impact not only on primary producer's biomass but also on the composition of phytoplankton communities by influencing the species that locally coexist (Cardinale et al., 2009a, 2009b and this study). In previous studies, catchment characteristics were shown to be good predictors of lakes' trophic status in the Paris area (Catherine et al., 2010, 2013). In addition, Soininen and Luoto (2012) recently suggested that catchment characteristics significantly impact lakes phytoplankton richness through their impact on water quality.

Here, we showed that catchments characteristics and meteorological data could be used to predict phytoplankton richness (Table 3). We found SR to be higher in water bodies connected to the hydrographic network and in catchment having a higher drainage index, both variables modulating the catchment's ability to transport nutrients (Young et al., 1996; Bremigan et al., 2008; Fraterrigo and Downing, 2008). In addition, we found that higher rainfall (RainJ20) had a positive impact on phytoplankton richness. Finally, a high percentage of forested cover in the catchment, which is inversely related to the main nutrient sources (i.e. urban and agricultural land covers; Catherine et al., 2010, 2013) reaching water bodies in the Paris area, was associated to lower phytoplankton species richness.

Using a variance partitioning approach, we showed that combining the information originating from the LS and CS models did not led to a significant increase in predictive power. The Random Forest approach is, by nature, not subjected to over-parameterization, as the variables selected to build tree are only included if they provide an increase in predictive power. Thus, this suggests that simple and easy to obtain variables, such as the one selected in the CS model, provide a meaningful way of predicting phytoplankton species richness in temperate lakes. Considering the link between phytoplankton richness and major ecosystem processes (e.g. resilience, productivity) and phenomena (e.g. occurrence of bloom-forming cyanobacteria), this approach may prove useful and cost-effective for the management and conservation of aquatic ecosystems.

5. Conclusions

Based on the analysis of phytoplankton species richness in a set of 50 peri-urban lakes, that seasonal temperature variations and resource availability were found to strongly modulate local phytoplankton species richness. Combining the two predictive models (based either on local-scale or catchment-scale variables) did not increase predictive accuracy; therefore suggesting that the catchment-scale model probably explains observed species richness variations through the impact of catchment-scale variables on in-lake water quality characteristics. We also found that both models overestimate observed species richness in winter and that some unexplained process probably drives richness down in winter.

Overall, this study provides an effective way of predicting phytoplankton species richness in temperate lakes that may be used for aquatic ecosystem management and conservation.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.scitotenv.2016.03.179>.

Autorship

SM, AC, CB and MT designed the study; SM and AC contributed data; SM, AC and DM contributed to data analysis; SM and AC wrote the manuscript; all authors contributed to editing the manuscript.

Acknowledgements

The authors acknowledge funding from the French National Research Agency (ANR, www.agence-nationale-recherche.fr) through the CYANOTOX (ANR-007-SEST-05) and PULSE (ANR-10-CEPL-0010) projects. This work was also supported by the CNRS through a PhD grant awarded to S. Maloufi (PED grant for students originating from developing countries). We thank the two anonymous reviewers for their useful comments on the earlier draft of the manuscript.

References

- Angeler, D.G., Drakare, S., 2013. Tracing alpha, beta, and gamma diversity responses to environmental change in boreal lakes. *Oecologia* 172, 1191–1202.
- Azeria, E.T., Fortin, D., Hébert, C., Peres-Neto, P., Pothier, D., Ruel, J.-C., 2009. Using null model analysis of species co-occurrences to deconstruct biodiversity patterns and select indicator species. *Divers. Distrib.* 15, 958–971.
- Bell, T., Newman, J.A., Silverman, B.W., Turner, S.L., Lilley, A.K., 2005. The contribution of species richness and composition to bacterial services. *Nature* 436, 1157–1160.
- Bellwood, D.R., Hughes, T.P., 2001. Regional-scale assembly rules and biodiversity of coral reefs. *Science* 292, 1532–1535.
- Biau, G., 2012. Analysis of a random forests model. *J. Mach. Learn. Res.* 13, 1063–1095.
- Bissinger, J.E., Montagnes, D.J.S., Sharples, J., Atkinson, D., 2008. Predicting marine phytoplankton maximum growth rates from temperature: improving on the Eppley curve using quantile regression. *Limnol. Oceanogr.* 53, 487–493.
- Borcard, D., Legendre, P., 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Model.* 153, 51–68.
- Borcard, D., Legendre, P., Avois-Jacquet, C., Tuomisto, H., 2004. Dissecting the spatial structure of ecological data at multiple spatial scales. *Ecology* 85, 1826–1832.
- Borcard, D., Gillet, F., Legendre, P., 2011. *Numerical Ecology With R*. Springer, New York.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Bremigan, B.T., Soranno, P.A., Gonzalez, M.J., Bunnell, D.B., Arend, K.K., Renwick, W.H., Stein, R.A., Vanni, M.J., 2008. Hydrogeomorphic features mediate the effects of land use/cover on reservoir productivity and food webs. *Limnol. Oceanogr.* 53, 1420–1433.
- Butterwick, C., Heaney, S.I., Talling, J.F., 2005. Diversity in the influence of temperature on the growth rates of freshwater algae, and its ecological relevance: temperature and growth rates of planktonic algae. *Freshw. Biol.* 50, 291–300.
- Cardinale, B.J., Bennett, D.M., Nelson, C.E., Gross, K., 2009a. Does productivity drive diversity or vice versa? A test of the multivariate productivity–diversity hypothesis in streams. *Ecology* 90, 1227–1241.
- Cardinale, B.J., Hillebrand, H., Harpole, W.S., Gross, K., Ptacnik, R., 2009b. Separating the influence of resource “availability” from resource “imbalance” on productivity–diversity relationships. *Ecol. Lett.* 12, 475–487.
- Catherine, A., Troussellier, M., Bernard, C., 2008. Design and application of a stratified sampling strategy to study the regional distribution of cyanobacteria (Ile-de-France, France). *Water Res.* 42, 4989–5001.
- Catherine, A., Mouillot, D., Escoffier, N., Bernard, C., Troussellier, M., 2010. Cost effective prediction of the eutrophication status of lakes and reservoirs. *Freshw. Biol.* 55, 2425–2435.
- Catherine, A., Escoffier, N., Belhocine, A., Nasri, A.B., Hamlaoui, S., Yéprémian, C., Bernard, C., Troussellier, M., 2012. On the use of the FluoroProbe®, a phytoplankton quantification method based on fluorescence excitation spectra for large-scale surveys of lakes and reservoirs. *Water Res.* 46, 1771–1784.
- Catherine, A., Mouillot, D., Maloufi, S., Troussellier, M., Bernard, C., 2013. Projecting the impact of regional land-use change and water management policies on lake water quality: an application to periurban lakes and reservoirs. *PLoS ONE* 8, e72227.
- Chase, J.M., Leibold, M.A., 2002. Spatial scale dictates the productivity–biodiversity relationship. *Nature* 416, 427–430.
- Clarke, A., Gaston, K.J., 2006. Climate, energy and diversity. *Proc. R. Soc. B Biol. Sci.* 273, 2257–2266.
- Connell, J.H., 1978. Diversity in tropical rain forests and coral reefs. *Science* 199, 1302–1310.
- Currie, D.J., 1991. Energy and large-scale patterns of animal- and plant-species richness. *Am. Nat.* 137, 27–49.
- Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guegan, J.-F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E., Turner, J.R.G., 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* 7, 1121–1134.
- de Barros, C.F.A., dos Santos, A.M.M., Barbosa, F.A.R., 2013. Phytoplankton diversity in the middle Rio Doce lake system of southeastern Brazil. *Acta Bot. Bras.* 27, 327–346.
- de Nogueira, I.S., Nabout, J.C., do Ibañez, M.S.R., Bourgoin, L.M., 2010. Determinants of beta diversity: the relative importance of environmental and spatial processes in structuring phytoplankton communities in an Amazonian floodplain. *Acta Limnol. Brasiliensis* 22, 247–256.
- Dodson, S.I., Arnett, S.E., Cottingham, K.L., 2000. The relationship in lake communities between primary productivity and species richness. *Ecology* 81, 2662–2679.
- Duffy, J.E., 2009. Why biodiversity is important to the functioning of real-world ecosystems. *Front. Ecol. Environ.* 7, 437–444.
- Eaton, A.D., Clesceri, L.S., Rice, E.W., Greenberg, A.E., Franson, M.A.H., 2005. *Standard Methods for the Examination of Water and Wastewater: Centennial Edition*. American Public Health Association, Washington DC, USA.
- Eloranta, P., 1993. Diversity and succession of the phytoplankton in a small lake over a two-year period. *Hydrobiologia* 249, 25–32.
- EN 15204:2006, 2006. *Water Quality — Guidance Standard on the Enumeration of Phytoplankton Using Inverted Microscope (Utermöhl Technique)*. European Committee for Standardization, Brussels.
- Evans, J.S., Murphy, M.A., 2015. rUtilities: Random forests model selection and performance evaluation.
- Filstrup, C.T., Hillebrand, H., Heathcote, A.J., Harpole, W.S., Downing, J.A., 2014. Cyanobacteria dominance influences resource use efficiency and community turnover in phytoplankton and zooplankton communities. *Ecol. Lett.* 17, 464–474.
- Foissner, W., 2006. Biogeography and dispersal of micro-organisms: a review emphasizing protists. *Acta Protozool.* 45, 111–136.
- Fraterrigo, J.M., Downing, J.A., 2008. The influence of land use on lake nutrients varies with watershed transport capacity. *Ecosystems* 11, 1021–1034.
- Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M.V., Green, J.L., Brown, J.H., 2008. A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci.* 105, 7774–7778.
- Gillman, L.N., Wright, S.D., 2006. The influence of productivity on the species richness of plants: a critical assessment. *Ecology* 87, 1234–1243.
- Green, J., Bohannan, B.J.M., 2006. Spatial scaling of microbial biodiversity. *Trends Ecol. Evol.* 21, 501–507.
- Groner, E., Novoplansky, A., 2003. Reconsidering diversity–productivity relationships: directness of productivity estimates matters. *Ecol. Lett.* 6, 695–699.
- Horner-Devine, M.C., Lage, M., Hughes, J.B., Bohannan, B.J.M., 2004. A taxa–area relationship for bacteria. *Nature* 432, 750–753.
- IAURIF (Institute for Urban Planning of the Ile-de-France Region), 2014R. *Key Figures of the Ile-de-France Region*.
- Interlandi, S.J., Kilham, S.S., 2001. Limiting resources and the regulation of diversity in phytoplankton communities. *Ecology* 82, 1270–1282.
- Irigoin, X., Huisman, J., Harris, R.P., 2004. Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature* 429, 863–867.
- Ishwaran, H., Kogalur, U.B., 2014. randomForestSRC: Random Forests for Survival, Regression and Classification (RF-SRC).
- Janitz, S., Strobl, C., Boulesteix, A.-N., 2013. An AUC-based permutation variable importance measure for random forests. *BMC Bioinform.* 14, 119.
- Kühn, I., 2007. Incorporating spatial autocorrelation may invert observed patterns. *Divers. Distrib.* 13, 66–69.
- Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659.
- Legendre, P., Legendre, L., 2012. *Numerical Ecology*. Third English ed. Elsevier, Amsterdam.
- Leibold, M.A., 1996. A graphical model of keystone predators in food webs: trophic regulation of abundance, incidence, and diversity patterns in communities. *Am. Nat.* 147, 784–812.
- Levin, S.A., 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology* 73, 1943.
- Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R News* 2, 18–22.
- Loreau, M., Naeem, S., Inchausti, P. (Eds.), 2002. *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*. Oxford University Press, Oxford.
- Maloufi, S., Catherine, A., Mouillot, D., Louvard, C., Couté, A., Bernard, C., Troussellier, M., 2016. Environmental heterogeneity among lakes promotes hyper β -diversity across phytoplankton communities. *Freshw. Biol.* <http://dx.doi.org/10.1111/fwb.12731> (in press).
- Marquet, P.A., Fernández, M., Navarrete, S.A., Valdovinos, C., 2004. Diversity emerging: toward a deconstruction of biodiversity patterns. In: Lomolino, M.V., Heaney, L.R. (Eds.), *Frontiers of Biogeography: New Directions in the Geography of Nature*. Sinauer Associates, Sunderland, USA, pp. 191–209.
- Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V.H., Staley, J.T., 2006. Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* 4, 102–112.
- Mercier, L., Darnaud, A.M., Brugier, O., Vasconcelos, R.P., Cabra, H.N., Costa, M.J., Lara, M., Jones, D.L., Mouillot, D., 2011. Selecting statistical models and variable combinations for optimal classification using otolith microchemistry. *Ecol. Appl.* 21, 1352–1364.
- Middelboe, A.L., Markager, S., 1997. Depth limits and minimum light requirements of freshwater macrophytes. *Freshw. Biol.* 37, 553–568.
- Millennium Ecosystem Assessment (MEA), 2005. *Current State and Trends Assessment*. Washington D.C., USA.
- Mittelbach, G.G., Steiner, C.F., Scheiner, S.M., Gross, K.L., Reynolds, H.L., Waide, R.B., Willig, M.R., Dodson, S.I., Gough, L., 2001. What is the observed relationship between species diversity and productivity? *Ecology* 82, 2381–2396.
- Murphy, M.A., Evans, J.S., Storer, A., 2010. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91, 252–261.
- Nabou, J.C., de Nogueira, I.S., de Oliveira, L.G., Morais, R.R., 2007. Phytoplankton diversity (alpha, beta, and gamma) from the Araguaia River tropical floodplain lakes (central Brazil). *Hydrobiologia* 575, 455–461.
- Özkan, K., Jeppesen, E., Søndergaard, M., Lauridsen, T.L., Liboriussen, L., Svenning, J.-C., 2013. Contrasting roles of water chemistry, lake morphology, land-use, climate and spatial processes in driving phytoplankton richness in the Danish landscape. *Hydrobiologia* 710, 173–187.
- Palmer, W.P., 1994. Variation in species richness: towards a unification of hypotheses. *Folia Geobot. Phytotaxon.* 29, 511–530.
- Ptácnik, R., Lepistö, L., Willén, E., Brettum, P., Andersen, T., Rekolainen, S., Lyche Solheim, A., Carvalho, L., 2008a. Quantitative responses of lake phytoplankton to eutrophication in Northern Europe. *Aquat. Ecol.* 42, 227–236.
- Ptácnik, R., Solimini, A.G., Andersen, T., Tamminen, T., Brettum, P., Lepistö, L., Willén, E., Rekolainen, S., 2008b. Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *Proc. Natl. Acad. Sci.* 105, 5134–5138.

- Ptácnik, R., Andersen, T., Brettum, P., Lepistö, L., Willén, E., 2010. Regional species pools control community saturation in lake phytoplankton. *Proc. R. Soc. B Biol. Sci.* 277, 3755–3764.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, R Foundation for Statistical Computing, Vienna Austria.
- Rao, C.R., 1964. The use and interpretation of principal component analysis in applied research. *Sankhya Ser. A* 26, 329–358.
- Reynolds, C.S., 2006. *Ecology of Phytoplankton*. Cambridge University Press, Cambridge, United Kingdom.
- Smith, V.H., 2007. Microbial diversity–productivity relationships in aquatic ecosystems: diversity–productivity relationships. *FEMS Microbiol. Ecol.* 62, 181–186.
- Smith, V.H., Foster, B.L., Grover, J.P., Holt, R.D., Leibold, M.A., deNoyelles, F., 2005. Phytoplankton species richness scales consistently from laboratory microcosms to the world's oceans. *Proc. Natl. Acad. Sci.* 102, 4393–4396.
- Soininen, J., Luoto, M., 2012. Is catchment productivity a useful predictor of taxa richness in lake plankton communities? *Ecol. Appl.* 22, 624–633.
- Stomp, M., Huisman, J., Mittelbach, G.G., Litchman, E., Klausmeier, C.A., 2011. Large-scale biodiversity patterns in freshwater phytoplankton. *Ecology* 92, 2096–2107.
- Streibel, M., Ptácnik, R., Stibor, H., Behl, S., Berninger, U., Haupt, F., Hingsamer, C., Mangold, C., Ptácnikova, R., Steinböck, M., Stockenreiter, M., Wickham, S., Wollrab, S., 2010. Water Column Stratification, Phytoplankton Diversity and Consequences for Resource Use and Productivity. Proceedings of the HYDRALAB III Joint User Meeting, Hannover, February 2010.
- Strong, D.R., 2010. Evidence and inference: shapes of species richness–productivity curves¹. *Ecology* 91, 2534–2535.
- Tibby, J., Reid, M.A., 2004. A model for inferring past conductivity in low salinity waters derived from Murray River (Australia) diatom plankton. *Mar. Freshw. Res.* 55, 597.
- Tilman, D., 1985. The resource ratio hypothesis of succession. *Am. Nat.* 125, 827–852.
- Turner, M.G., Gardner, R.H., O'Neill, R.V., 2001. *Landscape Ecology in Theory and Practice: Pattern and Process*. Springer, New York.
- Vallina, S.M., Follows, M.J., Dutkiewicz, S., Montoya, J.M., Cerneno, P., Loreau, M., 2014. Global relationship between phytoplankton diversity and productivity in the ocean. *Nat. Commun.* 5.
- Wetzel, R.G., 2001. *Limnology*. second ed. Saunders, Philadelphia.
- Wright, D.H., 1983. Species-energy theory: an extension of species-area theory. *Oikos* 41, 496–506.
- Young, W.J., Marston, F.M., Davis, R.J., 1996. Nutrient exports and land use in Australian catchments. *J. Environ. Manag.* 47, 165–183.