

ORIGINAL ARTICLE

Open Access



# Distribution patterns of epilithic diatoms along climatic, spatial and physicochemical variables in the Baltic Sea

Leena Virta<sup>1,2\*</sup> and Janne Soininen<sup>1</sup>

## Abstract

The species richness and community composition of the diatom communities were studied in the Baltic Sea, Northern Europe, to enhance knowledge about the diversity of these organisms in a brackish water ecosystem. Many organisms in the Baltic Sea have been studied extensively, but studies investigating littoral diatoms are scarce. The goal of this study was to examine the importance of climatic, spatial and water physicochemical variables as drivers of epilithic diatoms in the Gulf of Finland and the Gulf of Bothnia. The variation in species richness was best explained by pH, total phosphorus and total nitrogen. Redundancy Analysis indicated that the most important factors correlating with species composition were air temperature, silicon, total phosphorus, water temperature, salinity and pH. Variation Partitioning showed that the species composition was mostly affected by climatic and spatial variables, whereas physicochemical variables had little impact. However, the strongest factor was the combined influence of climatic, spatial and physicochemical variables. The results suggest that diatom species richness in the northern Baltic Sea is primarily regulated by local factors, while climatic and spatial variables have little impact on richness. Species composition is mostly affected by climatic and spatial variables. We conclude that understanding the distribution patterns of Baltic Sea diatoms requires the inclusion of climatic, spatial and water chemistry variables.

**Keywords:** Baltic Sea, Brackish, Diatoms, Epilithic, Salinity

## Introduction

Diatoms provide ca. 40% of the global marine primary production, and thus they play a fundamental role in food webs and chemical cycles in the aquatic ecosystems [1]. Changes in the diatom composition therefore affect entire ecosystems. The variation in diatom composition and species richness has been the subject of a large body of research both in marine and fresh waters [e.g. 2, 3]. Traditionally, diatoms are thought to respond merely to local environmental variables, such as water chemistry, and in many cases, water chemistry variables, such as nutrients (nitrogen, phosphorus and silicon), pH, salinity and conductivity have been found to be the most important explanatory variables for diatom composition

[e.g. 1, 2, 4]. However, the roles of climate and dispersal limitation have recently also been emphasized in explaining diatom distributions [3]. Organisms have distinct regional pools of species, and both climate and local physicochemical factors act as filters that determine the species composition in local communities [3, 5]. The relative importance of climatic factors and local physicochemical factors varies with study scale: climatic factors may override the local physicochemical factors at continental scales, while at regional scales (100–3000 km) diatom communities are influenced by both climatic and local factors [5].

The Baltic Sea is one of the world's three large brackish water seas. Its surface water salinity ranges from ca. 9 in the South-West to almost zero in the North-East [6]. Such a decrease in salinity is caused by the restricted exchange of water with the North Sea through the narrow Danish Straits. A number of rivers also discharge from the catchment to different parts of the Baltic

\*Correspondence: leena.virta@helsinki.fi

<sup>1</sup> Department of Geosciences and Geography, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland

Full list of author information is available at the end of the article

Sea—mostly to the Gulf of Bothnia, the Gulf of Finland and the Gulf of Riga. The largest annual total river runoff flows into the Gulf of Bothnia and the Gulf of Finland [7], which explains their low salinities. The Gulf of Bothnia and the Gulf of Finland also receive most of the incoming nutrient loads and organic material due to river runoffs. Anthropogenic nutrient inputs result in serious eutrophication including internal loading and hypoxia [8].

The Baltic Sea is a young formation with a low biodiversity [9]. Present conditions have only prevailed for ca. 3000 years, which is too short a period for brackish water communities to develop, and thus only a few freshwater or marine species have been able to adapt to live in the Baltic Sea [10]. Consequently, most important marine animal and algae groups are missing. This situation makes the whole ecosystem sensitive to disturbances as the extinction of a keystone species may affect the entire ecosystem [9].

Although diatoms are highly useful in indicating various environmental changes and they are used as standards in freshwater biomonitoring, diatom research has been somewhat scarce in the Baltic Sea. However, there have been important studies of benthic diatoms by Leskinen and Hållfors in the southernmost Finland [e.g. 11–13], Vilbaste et al. [14, 15] in the Gulf of Riga and the coastal area of Estonia, Sommer in the coastal area of Germany [16], and Snoeijs, Busse and Ulanova in the coastal area of Sweden [e.g. 17, 18]. However, sampling has usually covered only a limited spatial scale with the exception of the study by Ulanova et al. [19]. Thus, there is a lack of broad research studies investigating the effects of climatic, spatial and water physicochemical variables on benthic diatoms in the Baltic Sea.

The aim of this study was to determine the most important climatic, spatial and physicochemical variables that explain variation in diatom species richness and composition in the Baltic Sea, using epilithic (i.e. growing on stones) diatom data collected in the littoral zone of the Gulf of Finland and the Gulf of Bothnia. Sampling covered the entire Finnish coastline, and explanatory variables were represented by climatic (July air temperature and precipitation), spatial (latitude, longitude and coast line exposition) and water physicochemical (total nitrogen, total phosphorus, pH, salinity, dissolved oxygen, silicon and water temperature) variables.

## Materials and methods

### Study area

The study comprised 37 sampling sites located in the southern and western coasts of Finland, in the Gulf of Finland and the Archipelago Sea (sampling sites 1–13), the Bothnian Sea (sampling sites 14–23) and the Bothnian Bay (sampling sites 24–37). The northernmost

sampling site was in Kemi (65° 44' N, 24° 34' E), the southernmost in Hanko (59° 49' N, 22° 58' E), the westernmost in Kustavi (60° 33' N, 21° 22' E) and the easternmost in Virolahti (60° 35' N, 27° 42' E) (Fig. 1).

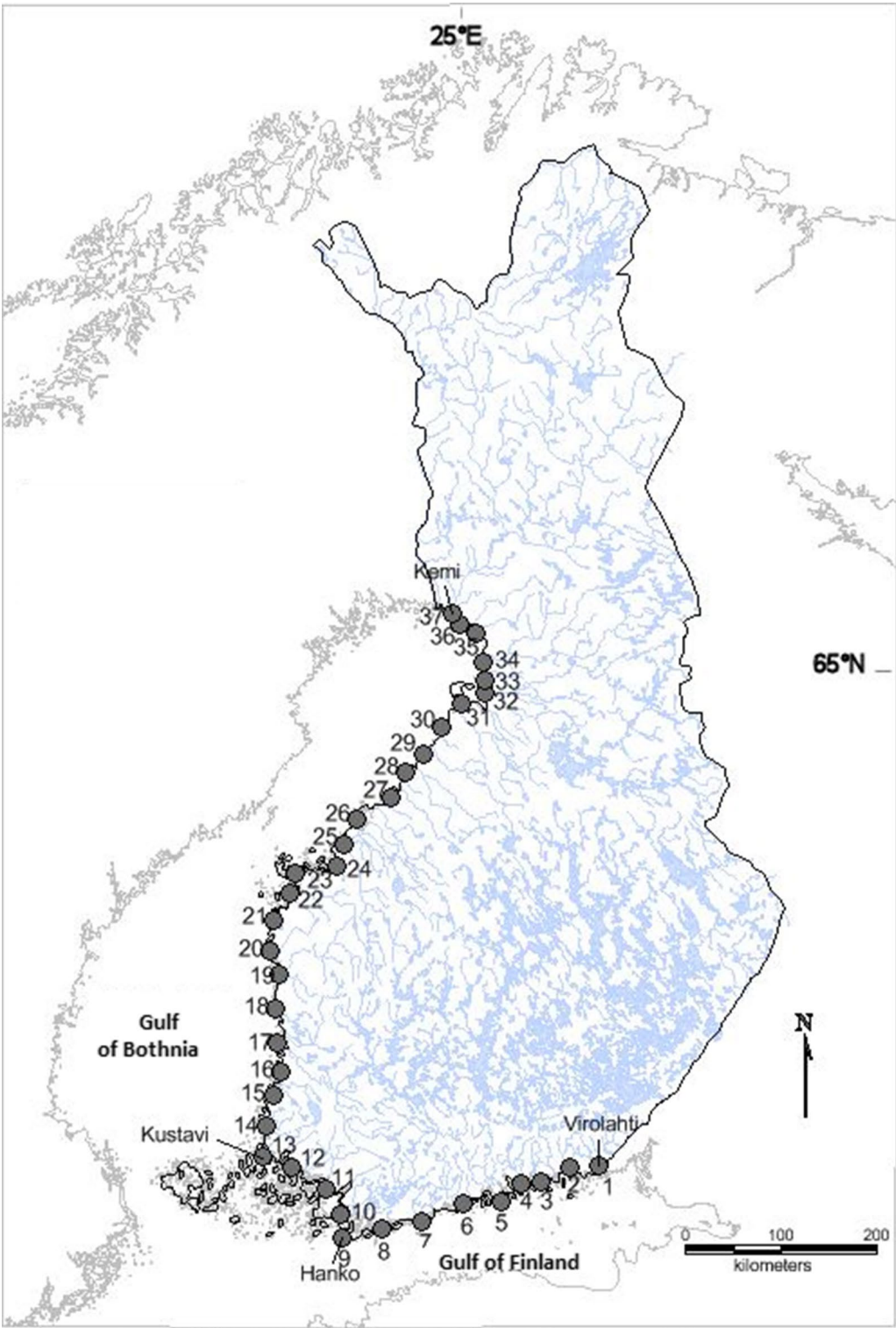
The Gulf of Finland is the easternmost part of the Baltic Sea and is bounded by Finland, Estonia and Russia. Unlike several other parts of the Baltic Sea, the western outline of the Gulf of Finland does not form a hydrographic threshold [7]. The Finnish archipelago in the Gulf of Finland is shallow, disintegrated and contains many islands. The Neva River in the eastern end of the Gulf of Finland produces the largest single fresh water inflow into the Baltic Sea [20]. The Bothnian Sea and the Bothnian Bay belong to the Gulf of Bothnia, which is a northern extension of the Baltic Sea. The Gulf of Bothnia is isolated from the rest of the Baltic Sea by an archipelago and a hydrographic threshold, which leads to substantially different environmental features [21]. River discharges lower the salinity in the Gulf of Bothnia (ca. 0 in the North and ca. 6 in the South) and the Gulf of Finland (<2 in the East and ca. 6 in the West) [20, 21] and bring most of the nutrient loads thus increasing the nutrient levels especially in the littoral zone [22]. Both gulfs suffer from eutrophication and water quality deterioration, but the issue is even more severe in the Gulf of Finland.

Biogeographically, the studied area belongs to the boreal ecoregion. The climate represents an ecotone between continental and oceanic climate zones [23]. The variation in annual mean temperatures is from 1.6 °C (Kemi) to 6.0 °C (Hanko), and the annual precipitation varies between 580 mm and 768 mm [24].

### Diatom sampling and laboratory analysis

The diatom samples were collected between the 1st and 15th of July 2013 in sheltered bays to minimize the effect of waves. The distance between sampling sites was ca. 30 km (Fig. 1), but large river mouths with ample nutrient loads were avoided so that minimum distance to river mouths was kept to 1 km. The diatom sampling complied with a standard SFS-EN 13946 adapted by Eloranta et al. [25]. Five stones with a diameter of at least 10 cm were randomly collected from each site from depths of 20–50 cm, and diatoms were sampled with a toothbrush from a 25 cm<sup>2</sup> surface area of each stone. The samples were preserved with ethanol in the field.

The samples were processed according to the SFS-EN 13946 standard. Organic material was removed from the samples by boiling with hydrogen peroxide (30% H<sub>2</sub>O<sub>2</sub>). Cleaned diatoms were mounted on slides using Naphrax [25]. To ensure an adequate number of counted valves per sample, two slides were prepared for each sample. Approximately 300 valves per sample were counted and identified to species level (when possible) with a light



**Fig. 1** Map of the study area, showing the 37 sampling sites, including the northernmost (Kemi), westernmost (Kustavi), southernmost (Hanko) and easternmost (Virolahti) sites

microscope (magnification 1000 $\times$ ) according to Krammer and Lange-Bertalot [26]. The most recent names of the taxa were verified using AlgaeBase [27]. The counting of 300 valves per sample was documented to be an adequate number for a reliable diatom community analysis in most environments [28].

#### Water sampling and laboratory analyses

Water samples (0.5 l) were collected simultaneously with diatom sampling and they were used for the analysis of total phosphorus complying to standard SFS-EN 1189. Water temperature and conductivity were measured in situ using Mettler Toledo MX300 probe. Conductivity was used to calculate the salinity complying the UNESCO formula [29].

In order to include a higher number of potentially important explanatory variables, we also collected environmental data from existing databases. The values of total nitrogen, pH, dissolved oxygen and silicon for sampling sites were interpolated from the materials of the Finnish Environment Institute (2013, raster, spatial resolution 20 m). The coast line exposition was also interpolated from the materials of the Finnish Environment Institute (2004, raster, spatial resolution 20 m). July air temperature and July precipitation were calculated using data from the Finnish Meteorological Institute (1981–2010, grid, spatial resolution 10 m) [24] with the Spatial Analyst tools of ArcGIS 10.2.1, Arcmap-applications [30]. For a list of all the explanatory variables, see Additional file 1.

#### Data analyses

Intercorrelations among the explanatory variables were assessed using Spearman's rank correlation coefficient ( $r_s$ ). Six of the correlations were diagnosed as strongly correlated ( $r_s < -0.7$  or  $> 0.7$ ,  $p < 0.001$ ) [31] (see Additional file 2 for details), and thus silicon, latitude, and air temperature were excluded from species richness analyses.

Correlograms with Moran's  $I$  [32] were constructed to examine the degree of spatial autocorrelation in the environmental variables and diatom species richness. The correlograms show correlation values in the range of  $[-1, 1]$ , where values near  $-1$  signify strong negative spatial autocorrelation, 0 complete randomness, and values near 1 strong positive spatial autocorrelation [33]. Autocorrelation coefficients were calculated for distance classes with 80 km intervals generating a spatial correlogram with 9 distance classes. The significance of the spatial autocorrelation in the distance classes was tested using the Bonferroni criterion  $\alpha = 0.05/k$ , where  $k$  is the number of distance classes used [34]. Thus, the corrected level of significance was set at  $\alpha = 0.05/9 = 0.00556$ .

Generalized Additive Models (GAM) were used to study the relationship between species richness and explanatory variables [35, 36]. GAMs are likelihood-based regression models that replace the linear function by an additive function [35]. Thus, they uncover the nonlinear covariate effects by estimating the shape of a smooth curve directly from the data [37]. GAMs were fitted using the R statistics package *mgcv* with maximum degrees of smoothing restricted to 3 [36]. The significance of the GAMs was analyzed using F-tests and Chi square tests [38].

Prior to statistical species composition analyses, water quality variables (except pH), climate variables and coast line exposition were log-transformed ( $\log_{10}(x)$ ) to reduce their skewed distributions, and diatom species composition data were Hellinger-transformed [39], because this method produces more precise estimates of the percentage of variation explained by the predictor variables [40].

As a constrained ordination, Redundancy Analysis (RDA) [41] was used to study the effects of climatic, spatial and water physicochemical variables on the diatom composition. Multicollinearity among the variables was assessed by determining variance inflation factors (VIF) [42], and values of  $VIF > 10$  were considered to represent high collinearity and thus removed from the analysis [43]. Due to high VIF, latitude was omitted from the RDA analysis.

Furthermore, variation partitioning was conducted to partition the variation in diatom community compositions with respect to all climatic (July air temperature and July precipitation), spatial (exposition, latitude and longitude) and water physicochemical variables (total phosphorus, total nitrogen, pH, salinity, oxygen, silicon and water temperature) and their combined effects [44]. Variation partitioning, based on and utilizing the eigenvalues of Redundancy Analysis, enables the determination of the individual and combined effects of local and regional explanatory variables and the proportion of unexplainable variation. All statistical analyses were conducted in the R environment [45] using packages *mgcv* [46], *ncf* [47], *rda* [48] and *vegan* [49].

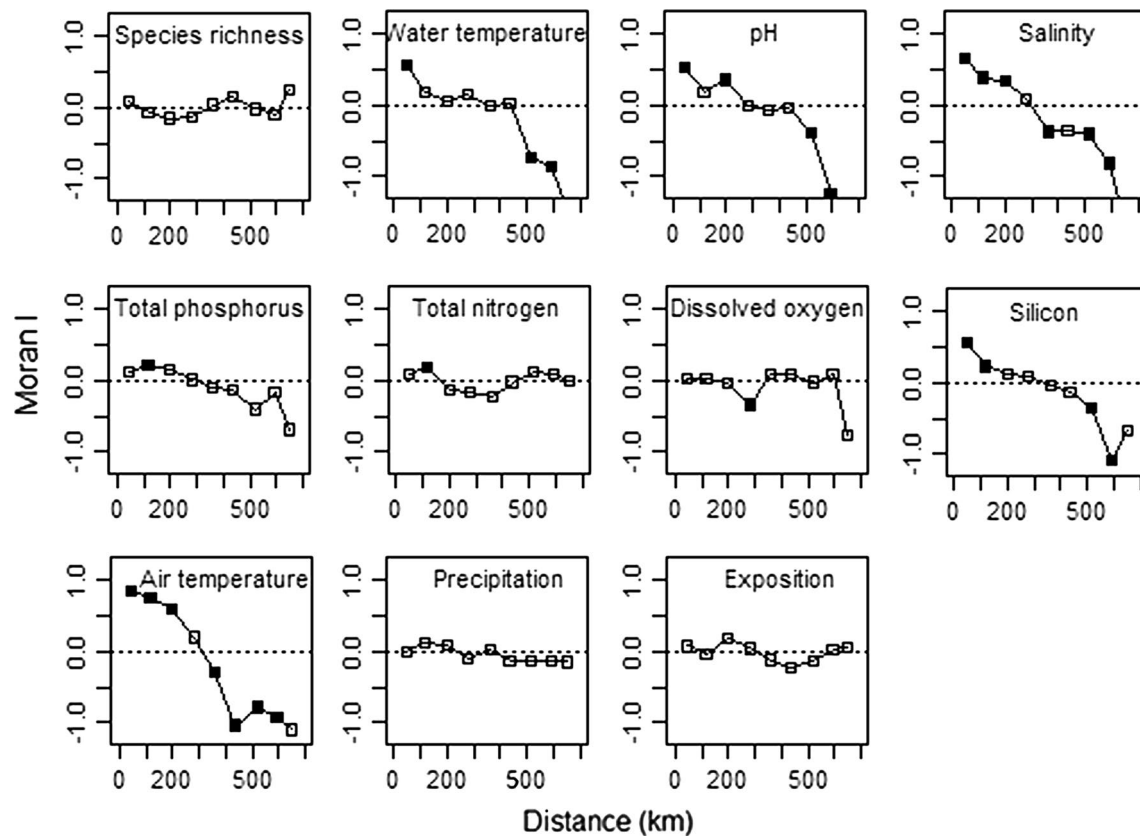
We then conducted Mantel tests to examine community turnover along geographical gradients [50]. The Mantel test is a method for modeling pairwise community dissimilarities as a function of pairwise spatial distances. Overall, the Mantel statistic  $r$  is a correlation between two dissimilarity or distance matrices.

## Results

#### Spatial autocorrelation in explanatory variables

Moran's correlograms for water temperature, pH, salinity, silicon and air temperature were significant in several distance classes according to the Bonferroni corrected





**Fig. 2** Spatial autocorrelation of environmental variables and species richness based on Moran's *I*. Filled squares indicate autocorrelation coefficients, which remain significant after Bonferroni correction at  $p = 0.05/9 = 0.0056$

level of significance ( $p < 0.00556$ ) (Fig. 2). Explanatory variables typically showed gradient-like structure, where correlation values were positive at short distances, but negative at longer distances. Importantly, no significant spatial autocorrelation was observed in the species richness data.

#### Characteristics of diatom assemblages

The number of species in a site varied between 16 and 54, and the mean richness was 34.0. A total of 230 diatom taxa belonging to 65 genera were identified. The genera with the highest number of species in the data were *Nitzschia* (35), *Navicula* (30) and *Achnanthes* (14). The most abundant species in the data were *Diatoma tenuis*, *Nitzschia frustulum* and *Fragilariforma virescens* var. *subsalina*. In addition, we observed 18 common taxa that represented at least 1.0% of the identified diatom cells, and whose average abundance in sampling sites also exceeded 1.0% (Table 1) (and see Additional file 3: Supplementary material for the entire list of diatom species). Some of the species were common (>1.0%

of cells) only in a particular region of the study area. For example, *Navicula incertata* was abundant only in the Gulf of Finland and the Archipelago Sea (sampling sites 2–13), while *Bacillaria paxillifera* (sampling sites 1–20), *Diatoma monoliformis* (1–25), *Fragilaria fasciculata* (2–21), *Navicula margalithii* (1–19), *Nitzschia frustulum* (1–24), *Nitzschia inconspicua* (3–26) and *Nitzschia liebetruithii* (2–23) were abundant in all areas except the Bothnian Bay. *Encyonema silesiacum* (28–37) and *Nitzschia fonticola* (24–37), were abundant only in the Bothnian Bay.

#### Species richness

Our best GAM model explained 35.2% of the total variation in species richness and indicated that diatom species richness had a statistically significant U-shaped relationship with pH ( $p < 0.01$ ) and a statistically significant positive relationship with total phosphorus ( $p < 0.01$ ) and total nitrogen ( $p < 0.05$ ) (Table 2; Fig. 3). All other variables had non-significant relationships with species richness.

**Table 1** Diatom taxa for which average abundance in sampling sites exceeded 1.0%. Mean and range of the species' relative abundance (%) and the percentage of sampling sites are also shown

	Abundance (%)	Sampling sites (%)
<i>Diatoma tenuis</i> Agardh	15.1 (0–56.7)	97.3
<i>Nitzschia frustulum</i> (Kützinger) Grunow	9.7 (0–74.0)	72.9
<i>Fragilariforma virescens</i> var. <i>subsalina</i> Grunow	8.7 (0–49.3)	94.6
<i>Achnanthes minutissimum</i> (Kützinger) Grunow	5.0 (0–64.0)	86.5
<i>Rhoicosphenia abbreviata</i> (Agardh) Lange-Bertalot	4.5 (0–34.0)	83.8
<i>Fragilaria capucina</i> Desmazières	4.5 (0–21.7)	91.9
<i>Nitzschia liebetruthii</i> Rabenhorst	4.0 (0–53.3)	45.9
<i>Staurosirella pinnata</i> (Ehrenberg) D.M.Williams & Round	3.7 (0–29.7)	75.7
<i>Fragilaria construens</i> (Ehrenberg) Grunow	3.3 (0–22.3)	73.0
<i>Achnanthes oblongella</i> Oestrup	2.8 (0–13.7)	91.9
<i>Fragilaria fasciculata</i> (Agardh) Lange-Bertalot	2.4 (0–23.3)	56.8
<i>Nitzschia perminuta</i> (Grunow) Peragallo	2.3 (0–11.3)	51.4
<i>Achnanthesiopsis delicatula</i> (Kützinger) Lange-Bertalot	2.3 (0–10.3)	86.5
<i>Epithemia sorex</i> Kützinger	2.2 (0–52.3)	43.2
<i>Nitzschia inconspicua</i> Grunow	2.1 (0–13.7)	62.2
<i>Navicula gregaria</i> Donkin	1.6 (0–5.7)	89.2
<i>Cocconeis placentula</i> var. <i>placentula</i> Ehrenberg	1.4 (0–6.0)	83.8
<i>Navicula incertata</i> Lange-Bertalot	1.4 (0–8.7)	32.4
<i>Navicula margalithii</i> Lange-Bertalot	1.1 (0–9.3)	43.2
<i>Halamphora coffeiformis</i> (Agardh) Levkov	1.1 (0–5.3)	70.3
<i>Nitzschia fonticola</i> (Grunow) Grunow	1.0 (0–15.0)	32.4

**Redundancy analysis and variation partitioning**

According to RDA, climatic, spatial and water phys-

**Table 2** Results of the best approximating Generalized Additive Model for explaining variation in diatom species richness

edf	Ref.df	Chi.sq	P value
Water temperature	1.693	1.904	4.903
pH	2.000	2.000	13.483
Salinity	1.000	1.000	1.229
Total phosphorus	1.788	1.952	9.732
Total nitrogen	1.675	1.889	7.922
Dissolved oxygen	1.628	1.853	3.452
Precipitation	1.000	1.000	1.117
Longitude	1.000	1.000	1.033
Exposition	1.000	1.000	0.891

Edf estimated degrees of freedom, Ref.df referenced degrees of freedom

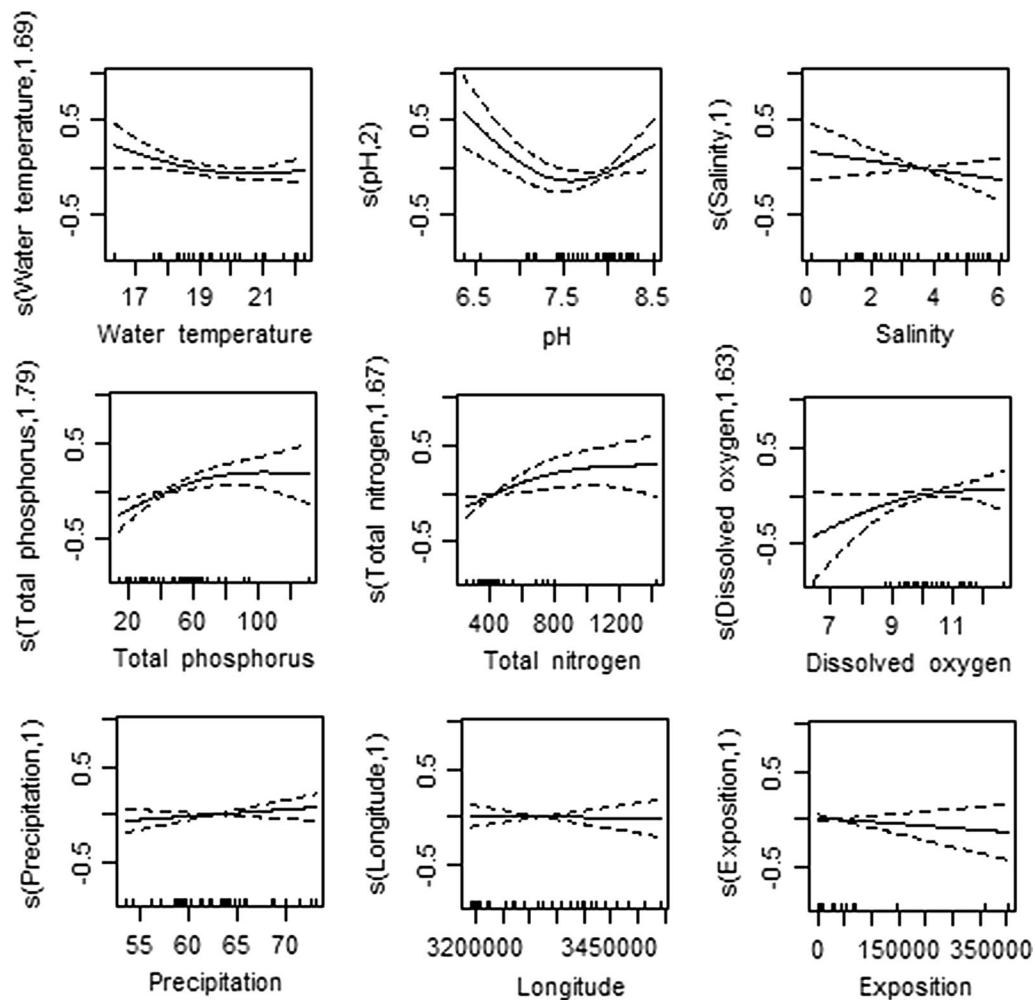
Statistical significances: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ;  $p < 0.1$

icochemical variables explained 43.3% of the variation in diatom community composition. The first RDA axis explained 17.7% of the variation and was mainly related to climate variable air temperature and water

physicochemical variables water temperature, silicon, total phosphorus, salinity and pH. This axis separated sampling sites in the Gulf of Finland and the Archipelago Sea (dominant species *Rhoicosphenia abbreviata*, *Nitzschia frustulum* and *Nitzschia liebetruthii*) on the left side of the diagram from the sampling sites in the Bothnian Bay (dominant species *Diatoma tenuis* and different species of *Fragilaria*) on the right side of the diagram (Fig. 4). The sampling sites in the Bothnian Sea were located in the central part of the diagram. The second axis explained 7.5% of variation and was mostly related to dissolved oxygen. No clear division among the study regions was observed on the second axis.

In variation partitioning of community composition, the combined effect of all explaining variables (8.8%) and the combined effect of spatial and water physicochemical variables (5.0%) accounted for most of the variation (Fig. 5). The individual effects of climatic, spatial and water physicochemical variables, as well as the combined effect of climatic and water physicochemical variables, and combined effect of climatic and spatial variables were low.

The Mantel test indicated that the community dissimilarity of diatoms was significantly correlated with geographical distance (see Additional file 4 for details).



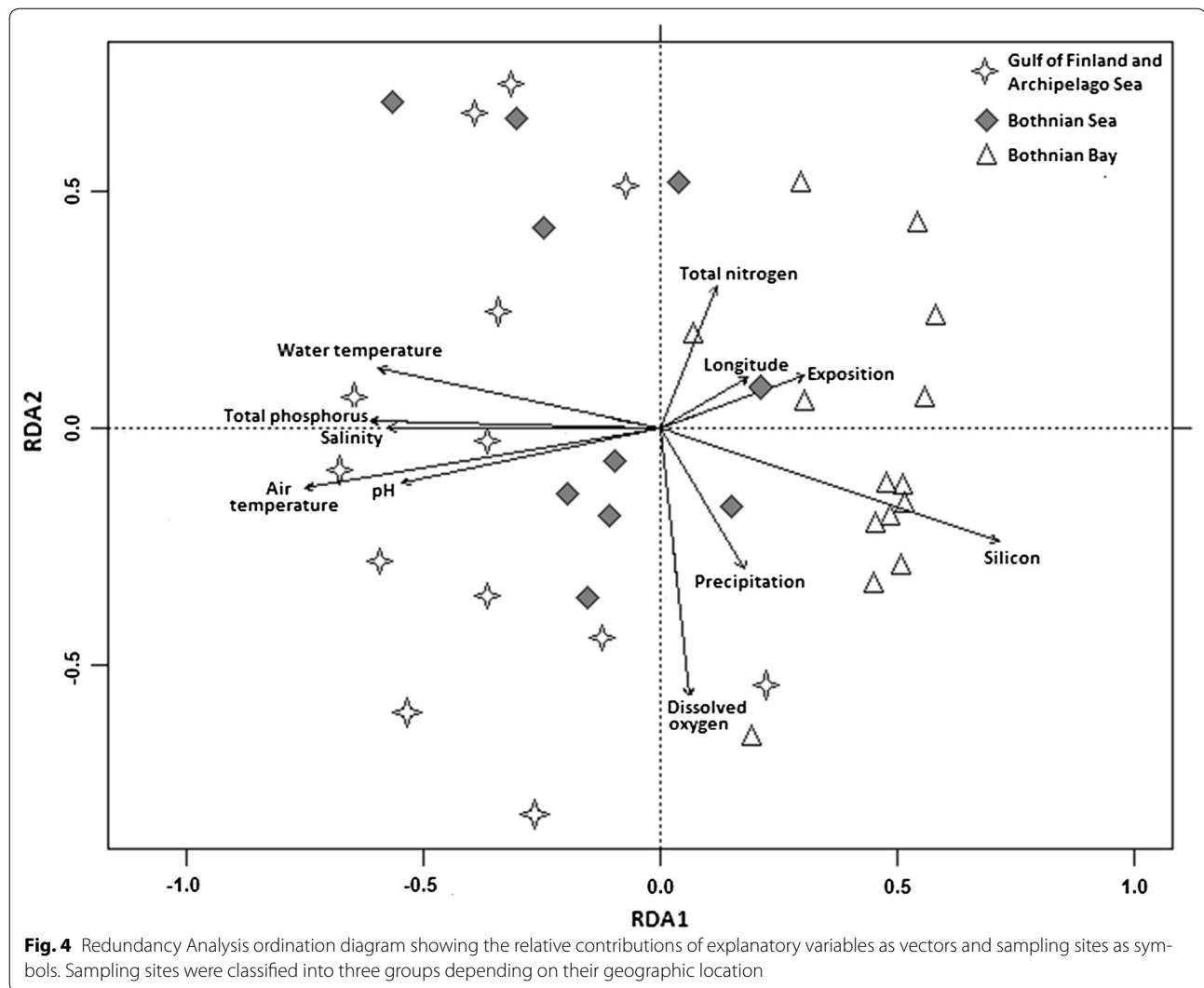
**Fig. 3** The results of best Generalized Additive Model explaining variation in diatom species richness. The dashed lines are the 95% of confidence intervals. On the y-axis, penalized regression splines (s) of each variable present their respective degrees of freedom

## Discussion

Our analyses showed that diatom species richness was related to three water chemistry variables: pH, total phosphorus and total nitrogen. The importance of pH on the diatom species richness in the Baltic Sea is in accordance with several earlier studies [e.g. 51]. However, our result of a U-shaped relationship between species richness and pH disagrees with some other studies, which have shown a unimodal relationship [e.g. 52]. We speculate that our different result may be due to some unmeasured environmental variables, such as trace metals or organic solvents. We also expect that covering a wider pH range would have further emphasized the importance of pH in regulating the species richness.

Previous studies investigating the effect of nutrients on diatom and phytoplankton species richness have settled upon contradictory results, and concluded that the

role of nutrients varies between regions [53]. Marine environments are usually considered as nitrogen limited [54], while freshwaters are often thought as phosphorus limited [55]. Previous studies have found differences in nutrient limitation between the Baltic Sea basins: the Bothnian Bay in the North seems to be phosphorus limited throughout the year, and the Gulf of Finland seems to be nitrogen limited, with the exception of some high nitrogen peaks in autumn. The nutrient limitation of the Bothnian Sea varies within the year [56] (see also [57] about joint limitation by N and P). Thus, we assumed that both total phosphorus and total nitrogen would be among the most important explanatory variables for diatom richness in our study [e.g. 11, 17]. However, we found that the effect of phosphorus on richness was slightly stronger than the nitrogen effect, which agrees with the findings of Blomqvist et al. [58]. This finding



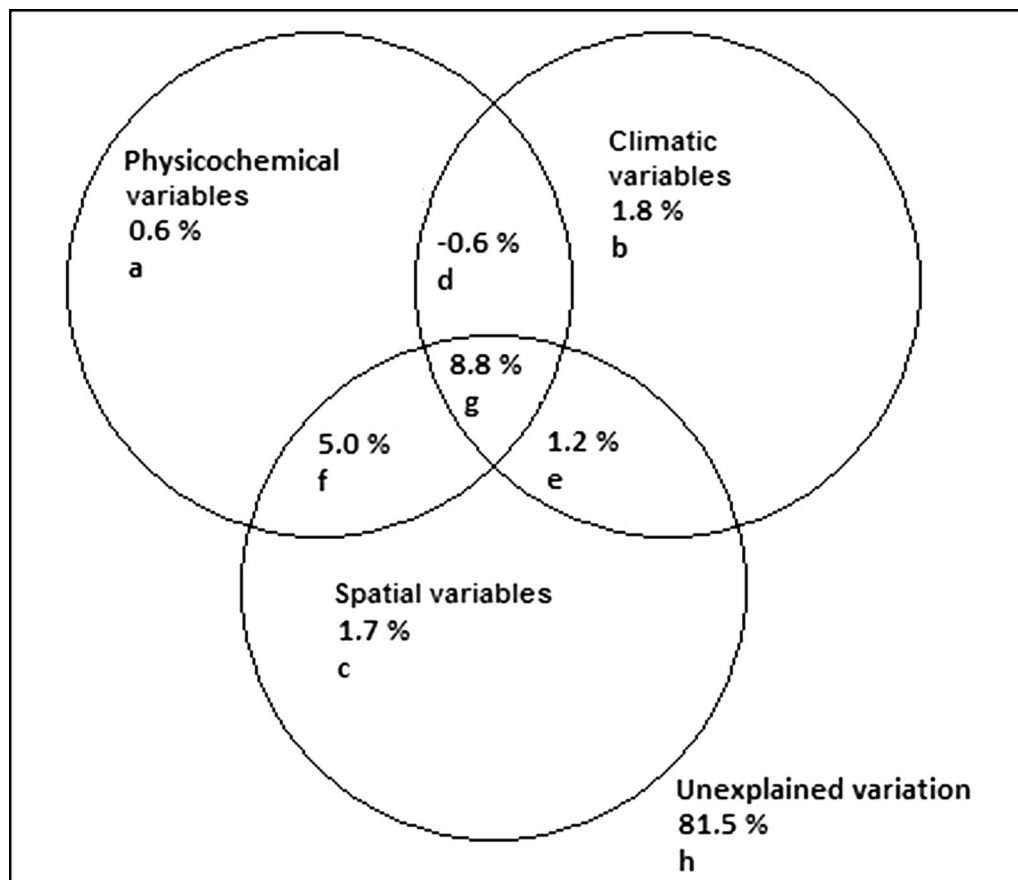
may reflect the fact that salinity in our study area was closer to freshwater than marine. Some earlier studies have observed a unimodal relationship between nutrients and diatom species richness, stemming from P limitation in low nutrient levels leading to increased P competition, and high pH in high P values [53, 59]. The nutrient levels in our sampling sites remained moderate, thus leading to positive, however leveling off, dependence between total phosphorus and species richness, and total nitrogen and species richness.

RDA showed that salinity was one of the key variables affecting diatom composition on the first axis, albeit its effect was somewhat weaker than air and water temperature, silicon and total phosphorus. The salinity gradient in our study area was 0.1–6.1, thus making the RDA result consistent with previous studies. Busse and Snoeijs [17], Ulanova and Snoeijs [18], and Ulanova et al. [19] studied the Swedish coast (salinity gradient approx. 0.4–11.4) and

emphasized the role of salinity as one of the most important explanatory variables for diatoms. Clarke et al. [60] studied the Danish Straits (salinity gradient 2.7–31.1) and noted a similar strong impact of salinity on diatoms. In contrast, Weckström and Juggins [59] (salinity gradient 0.7–6.4), and Leskinen and Hällfors [12] (salinity ca. 6) studied the Gulf of Finland and reported salinity to have only a minor effect on diatom communities. In conclusion, these studies show how both the extent of the salinity gradient and study area seem to affect the importance of salinity on diatoms.

RDA also showed that water and air temperatures were important variables for diatom compositions. Temperature may affect diatoms directly by influencing metabolism, growth and reproduction [61], or indirectly by changing the physical, chemical and biological characteristics of the water. Therefore, some earlier studies have noted air and water temperatures to be influential to





**Fig. 5** Variation partitioning of diatom communities using water physicochemical variables, climate and geographic location as explaining variables. *a–c* show the individual effects of variables; *d–g* show the combined effects of variables; *h* shows the amount of explained variation

diatoms [e.g. 62]. We also showed silicon to exert a strong effect on littoral diatom composition, which is inconsistent with several previous studies suggesting that silicon is a limiting nutrient only in the pelagic areas [12, 19].

Due to these driving variables, RDA clearly separated the southern and northern study areas based on their diatom communities. The southern sampling sites in the Gulf of Finland and the Archipelago Sea were characterized by high temperatures, salinity, pH and concentrations of total phosphorus, but low concentrations of silicon. Diatom samples collected from the southern area were dominated e.g. by *Navicula incertata*, which is a brackish water species tolerating somewhat high salinities, and *Nitzschia frustulum*, *Nitzschia liebetruhtii* and *Diatoma moniliformis*, which thrive in eutrophic waters and indicate high nutrient levels. Samples from the Bothnian Bay in the North, were dominated by the freshwater species *Encyonema silesiacum* and *Nitzschia fonticola*. Such regional segregation of species is typical for diatom studies [e.g. 63] due to climatic variation, restricted dispersal and regional differences in water chemistry.

According to the traditional view, unicellular organisms are ubiquitous and their community compositions are only regulated by local water chemistry variables [e.g. 64]. During the past decades, however, the traditional view has been questioned, and particularly the seminal approach of variation partitioning by Borcard et al. [44] has given rise to research concerning the role of local and regional variables affecting diatom community composition [e.g. 63, 65]. In particular, benthic algae have been documented to exhibit spatially more restricted distributions than previously believed, possibly because they are less exposed to wind and currents than planktonic algae [66]. In our variation partitioning, only 0.6% of variation in community composition was explained solely by local water physicochemical variables, whereas the spatial variables explained purely 1.7% of the variation and the combined effect of all explaining variables was 8.8%. Such a spatial component was also evident in Mantel test in which community dissimilarity was correlated with geographical distance. This may reflect either dispersal limitation or spatially structured environmental

variation. For example, Heino et al. [67] and Potapova and Charles [65] also reported that spatial variables had stronger influence on communities than water chemistry variables in boreal and North American streams, respectively. The importance of spatial variables on diatom community composition has been criticized, however, because spatial effect may also include the effects of any spatially structured, unmeasured environmental variables [67]. However, in our study, we included water chemistry variables that have been previously documented to be the most important drivers for diatom communities (e.g. pH, salinity, total nitrogen, total phosphorus) [e.g. 18]. Therefore, we assume that the importance of regional variables is not solely due to unmeasured water chemistry variables. In our study, climatic variables explained 1.8% of variation, i.e. slightly more than spatial variables. Previous studies taking into account the effect of climate on diatom composition are rare, but these studies have suggested that climate may be influential for diatoms [68]. Our study confirms that it is important for future diatom studies to consider climatic variables.

Our variation partitioning showed that 81.5% of variation remained unexplained. Such a high value is typical for corresponding studies [e.g. 67], and it may be due to the fact that we missed some abiotic (e.g. calcium and magnesium concentrations) or biotic (e.g. grazing pressure) factors that may have been influential to the community composition. While the individual effects of variable groups remained admittedly low here, the combined effects proved to be important. This result was to be expected, as spatial and climatic variables are strongly related, and both also affect water chemistry. In general, this shows that most of the variables act in concert to influence diatoms and any individual effect of a single variable is not easily disentangled.

## Conclusions

The usefulness of diatoms in indicating the local environment has been recognized, and diatoms are increasingly utilized in monitoring the state of the aquatic ecosystems. The Baltic Sea is a unique and vulnerable ecosystem with low salinity and biodiversity, and thus demands continuous research to support conservation efforts and monitoring programs. However, previous research on large-scale distribution patterns of diatoms in the Baltic Sea is scarce. Here, we showed that the most important variables affecting diatom species richness were pH, total phosphorus and total nitrogen. The most important variable groups affecting diatom composition were climatic and spatial variables, whereas the effect of water physicochemical variables was surprisingly weak. The combined effects of climatic, spatial and water physicochemical variables were, however, stronger. Therefore, we conclude that explanatory

variables affecting diatom species richness and composition are diverse, and understanding the distribution of diatoms requires the inclusion of not only the local water physicochemical variables, but also regional explanatory variables such as climatic and spatial variables.

## Additional files

**Additional file 1.** Explanatory variables.

**Additional file 2.** Statistical dependencies between explanatory variables.

**Additional file 3.** Diatom species list.

**Additional file 4.** Mantel test.

## Authors' contributions

LV and JS designed the study in cooperation. LV collected and analyzed the diatom and water samples, performed the data analyses, and was a major contributor in writing the manuscript. JS gave useful comments concerning the manuscript. Both authors read and approved the final manuscript.

## Author details

<sup>1</sup> Department of Geosciences and Geography, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland. <sup>2</sup> Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, 10900 Hanko, Finland.

## Acknowledgements

We thank the Walter and Andrée de Nottbeck Foundation for financial support. We are thankful for Hanna Reijola, Tuija Vaahtojärvi and Juhani Virkanen at the Environmental Laboratory of the Department of Geosciences and Geography (University of Helsinki) for the assistance with the analyses of water samples. We also thank Elina Virtanen (Finnish Environment Institute) and Miska Luoto (University of Helsinki) for providing us with useful data.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

The datasets generated and analyzed during the current study are available from the corresponding author on request, and will be available in ResearchGate after publication of the article.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Funding

Funding for the research has been received from the Walter and Andrée de Nottbeck Foundation.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 March 2017 Accepted: 7 September 2017

Published online: 19 September 2017

## References

1. Sanders R, Brown L, Henson S, Lucas M. New production in the Irminger Basin during 2002. *J Marine Syst.* 2005;55:291–310.

2. Sorokin YI, Sorokin PY. Production in the Sea of Okhotsk. *J Plankton Res*. 1999;21:201–30.
3. Pajunen V, Luoto M, Soininen J. Climate is an important driver for stream diatom distributions. *Glob Ecol Biogeogr*. 2016;25:198–206.
4. Hillebrand H, Sommer U. Diversity of benthic microalgae in response to colonization time and eutrophication. *Aquat Bot*. 2000;67:221–36.
5. Martiny JBH, Bohannan JM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Øvreås L, Reysenbach AL, Smith VH, Staley JT. Microbial biogeography: putting microorganisms on the map. *Microbiol Mol Biol R*. 2006;4:102–12.
6. Meier HEM, Kauker F. Sensitivity of the Baltic Sea salinity to the freshwater supply. *Clim Res*. 2003;24:231–42.
7. Leppäranta M, Myrberg K. *Physical Oceanography of the Baltic Sea*. Chichester: Springer; 2009.
8. Zander CD, Reimer LW. Parasitism at the ecosystem level in the Baltic Sea. *Parasitology*. 2002;124:119–35.
9. Johannesson K, Smolárz K, Grahn M, Andre C. The future of Baltic Sea populations: local extinction or evolutionary rescue? *Ambio*. 2011;40:179–90.
10. Elmgren R. Trophic dynamics in the enclosed, brackish Baltic Sea. *Rapp P-V Réun. Cons. Int. Explor. Mer*. 1984;183:152–69.
11. Leskinen E, Hällfors G. Community structure of epiphytic diatoms in relation to eutrophication on the Hanko Peninsula, south coast of Finland. 10th Diatom-Symposium 323–333; 1988.
12. Leskinen E, Hällfors G. Dynamics of epiphytic diatoms on *Cladophora glomerata* in the Tvärminne archipelago, northern Baltic Sea. in *Proceedings of the 11th International Diatom Symposium*; 1994.
13. Leskinen E, Hällfors G. *Tabularia waernii* (Diatomophyceae) in the northern Baltic Sea. *Ann Bot Fennici*. 1997;34:141–7.
14. Vilbaste S. The epiphytic and microphytobenthic diatoms in Estonian coastal waters (the Baltic Sea). in *Proceedings of the Thirteenth International Diatom Symposium*. 1995; 259–69.
15. Vilbaste S, Sundback K, Nilsson C, Truu J. Distribution of benthic diatoms in the littoral zone of the Gulf of Riga, the Baltic Sea. *Eur J Phycol*. 2000;35:373–85.
16. Sommer U. Nutrient competition experiments with periphyton from the Baltic Sea. *Mar Ecol Prog Ser*. 1996;140:161–7.
17. Busse S, Snoeijs P. Gradient responses of diatom communities in the Bothnian Sea (northern Baltic Sea), with emphasis on responses to water movement. *Phycologia*. 2003;42:451–64.
18. Ulanova A, Snoeijs P. Gradient responses of epilithic diatom communities in the Baltic Sea proper. *Estuar Coast Shelf Sci*. 2006;68:661–74.
19. Ulanova A, Busse S, Snoeijs P. Coastal diatom-environment relationships in the brackish Baltic Sea. *J Phycol*. 2009;45:54–68.
20. Andrejev O, Myrberg K, Alenius P, Lundberg PA. Mean circulation and water exchange in the Gulf of Finland: a study based on three-dimensional modelling. *Boreal Environ Res*. 2004;9:1–16.
21. Håkansson B, Alenius P, Brydsten L. Physical environment in the Gulf of Bothnia. *Ambio Special Report Number*. 1996;8: 5–12.
22. Cederswall H, Elmgren R. Biological effects of eutrophication in the Baltic Sea, particularly the coastal zone. *Ambio*. 1990;19:109–12.
23. Meier HEM. Baltic Sea climate in the late twenty-first century: a dynamical downscaling approach using two global models and two emission scenarios. *Clim Dyn*. 2006;27:39–68.
24. Pirinen P, Simola H, Aalto J, Kaukoranta J-P, Karlsson P, Ruuhela R. Tilastoja Suomen ilmastosta 1981–2010. Ilmatieteen laitoksen raportteja 2012: 1. Helsinki: Finnish Meteorological Institute.
25. Eloranta P, Karjalainen SM, Vuori KM. Piilevayhteisöt jokivesien ekologisen tilan luokittelussa ja seurannassa—menetelmäohjeet. Oulu: Painotupa ky; 2007.
26. Krammer K, Lange-Bertalot H. *Bacillariophyceae. Süßwasserflora von Mitteleuropa* 2 (1–4). Stuttgart: Gustav Fischer Verlag; 1986–1991.
27. Guiry MD, Guiry GM. *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. 2017. <http://www.algaebase.org>. Accessed 15 June 2017.
28. Teittinen A, Taka M, Ruth O, Soininen J. Variation in stream diatom communities in relation to water quality and catchment variables in boreal, urbanized region. *Sci Total Environ*. 2015;530–531:279–89.
29. UNESCO. Algorithms for computation of fundamental properties of seawater. UNESCO technical papers in marine science. 44; 1983.
30. ESRI. ArcGIS 10.2.1 for Desktop. Redlands, CA: Environmental System Research Institute; 2014.
31. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, García Márquez JR, Gruber B, Lafourcade B, Leitão PJ, Mündemüller T, McClean C, Osborne PE, Reineking J, Schröder B, Skidmore AK, Zurell D, Lautenbach S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. 2013;36:27–46.
32. Moran PAP. The interpretation of statistical maps. *J R Stat Soc B (Methodological)*. 1948;10:243–51.
33. Waldhör T. The spatial autocorrelation coefficient Moran's I under heteroscedasticity. *Stat Med*. 1996;15:887–92.
34. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995;310:170.
35. Hastie T, Tibshirani R. Generalized additive models. *Stat Sci*. 1986;1:297–318.
36. Guisan A, Edwards TC, Hastie T. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol Model*. 2002;157:89–100.
37. Benedetti A, Abrahamowicz M, Goldberg MS. Accounting for data-dependent degrees of freedom selection when testing the effect of a continuous covariate in generalized additive models. *Commun Stat Simul*. 2009;38(5):1115–35.
38. Wald A. A note on regression analysis. *Ann Math Stat*. 1947;18:586–9.
39. Legendre P, Gallagher ED. Ecologically meaningful transformations for ordination of species data. *Oecologia*. 2001;129:271–80.
40. Peres-Neto PR, Legendre P, Dray S, Borcard D. Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*. 2006;87:2614–25.
41. Van den Wollenberg AL. Redundancy analysis an alternative for canonical correlation analysis. *Psychometrika*. 1977;42:207–19.
42. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. *Qual Quant*. 2007;41:673–90.
43. Stine RA. Graphical interpretation of variance inflation factors. *Am Stat*. 1995;49:53–6.
44. Borcard D, Legendre P, Drapeau P. Partialling out the spatial component of ecological variation. *Ecology*. 1992;73:1045–55.
45. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. 2013. <http://www.R-project.org>. Accessed 21 Feb 2016.
46. Wood, S. Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation. 2016. <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>. Accessed 8 Feb 2016.
47. Bjørnstad ON. Spatial Nonparametric Covariance Functions. 2015. <https://cran.r-project.org/web/packages/ncf/ncf.pdf>. Accessed 8 Feb 2016.
48. Guo Y, Hastie T, Tibshirani R. Shrunken Centroids Regularized Discriminant Analysis. 2015. <https://cran.r-project.org/web/packages/rda/rda.pdf>. Accessed 8 Feb 2016.
49. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. *Vegan: Community ecology package*. 2016. <https://cran.r-project.org/web/packages/vegan/vegan.pdf>. Accessed 8 Feb 2016.
50. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res*. 1967;27:209–20.
51. Barinova S, Stenina A. Diatom diversity and ecological variables in the Arctic lakes of the Kostyanoi Nos Cape (Nenetsky Natural Reserve, Russian North). *Plant Biosyst*. 2013;2:397–410.
52. Soininen J. Responses of epilithic diatom communities to environmental gradients in some Finnish rivers. *Int Rev Hydrobiol*. 2002;87:11–24.
53. Soininen J. Is diatom diversity driven by productivity in boreal streams? *Diatom Res*. 2009;24:197–207.
54. Nixon SW, Ammerman JW, Atkinson LP, Berounsky VM, Billen G, Boicourt WC, Boynton WR, Church TM, Ditoro DM, Elmgren R, Garber JH, Giblin AE, Jahnke RA, Owens NJP, Pilson MEWQ, Seitzinger SP. The fate of nitrogen and phosphorus at the land-sea margin of the North Atlantic Ocean. *Biochemistry*. 1996;35:141–80.
55. Schindler DW. Eutrophication and recovery in experimental lakes: implications for lake management. *Science*. 1974;184(4139):897–9.
56. Danielsson Å, Papush L, Rahm L. Alterations in nutrient limitations—scenarios of a changing Baltic Sea. *J Marine Syst*. 2008;73:263–83.
57. Elser JJ, Bracken MES, Cleland EE, Gruner DS, Harpole WS, Hillebrand H, Ngai JT, Seabloom EW, Shurin JB, Smith JE. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol Lett*. 2007;10:1135–42.

58. Blomqvist S, Gunnars A, Elmgren R. Why the limiting nutrient differs between temperate coastal seas and freshwater lakes: a matter of salt. *Limnol Oceanogr*. 2004;49:2236–41.
59. Weckström K, Juggins S. Coastal diatom-environment relationships from the Gulf of Finland, Baltic Sea. *J Phycol*. 2005;42:21–35.
60. Clarke A, Juggins S, Conley D. A 150-year reconstruction of the history of coastal eutrophication in Roskilde Fjord, Denmark. *Mar Pollut Bull*. 2003;46:1615–8.
61. Bothwell ML. Growth-rate responses of lotic periphytic diatoms to experimental phosphorus enrichment: the influence of temperature and light. *Can J Fish Aquat Sci*. 1988;45:261–70.
62. Leland HV. Distribution of phytobenthos in the Yakima River basin, Washington, in relation to geology, land-use, and other environmental-factors. *Can J Fish Aquat Sci*. 1995;52:1108–29.
63. Soininen J, Paavola R, Muotka T. Benthic diatom communities in boreal streams: community structure in relation to environmental and spatial gradients. *Ecography*. 2004;27:330–42.
64. Finlay BJ. Global dispersal of free-living microbial eukaryote species. *Science*. 2002;296:1061–3.
65. Potapova MG, Charles DF. Benthic diatoms in USA rivers: distributions along spatial and environmental gradients. *J Biogeogr*. 2002;29:167–87.
66. Soininen J. Assessing the current related heterogeneity and diversity patterns of benthic diatom communities in a turbid and a clear water river. *Aquat Ecol*. 2004;38:495–501.
67. Heino J, Bin LM, Karjalainen SM, Mykrä H, Soininen J, Vieira LCG, Diniz JAF. Geographical patterns of micro-organismal community structure: are diatoms ubiquitously distributed across boreal streams? *Oikos*. 2010;119:129–37.
68. Clarke G, Kernan M, Marchetto A, Sorvari S, Catalan J. Using diatoms to assess geographical patterns of change in high-altitude European lakes from pre-industrial times to the present day. *Aquat Sci*. 2005;67:224–36.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

