*Article*

# Recognition of Patterns of Benthic Diatom Assemblages within a River System to Aid Bioassessment

**Shuxin Mao [1,2,3] , Shuhan Guo [1,3] , Hongbing Deng [2], Zhicai Xie [1] and Tao Tang [1,*]**

[1] State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China; sxmao_st@rcees.ac.cn (S.M.); guoshuhan@ihb.ac.cn (S.G.); zhcxie@ihb.ac.cn (Z.X.)
[2] State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China; denghb@rcees.ac.cn
[3] University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: tangtao@ihb.ac.cn; Tel.: +86-027-68780032

**Abstract:** Benthic algae, especially diatoms, are commonly used to assess water quality in rivers. However, algal-based assessments are challenging at the river system scale because longitudinal variation in physical habitat conditions may obscure algal responses to changes in water quality. In the present study, we surveyed benthic diatoms and environmental variables from a mountainous Chinese river system. Hierarchical clustering, discrimination analysis, and indicator species analysis were used together to explore associations between distribution patterns of diatom assemblages and water quality variables. Study sites were clustered into five groups based on their diatom community composition, with sites grouped by the sampling months. Chemical oxygen demand (COD), elevation, and total nitrogen (TN) were the most important predictors for site classification. Site groups with higher elevations had higher TN concentrations; however, COD concentrations were higher in lower elevation groups. Moreover, COD concentrations significantly differed between temporally separated groups. In total, 49 indicator species were identified for individual groups, with most taxa indicating the eutrophic condition. Additionally, we found that European diatom indices are not closely associated with water quality variables. We conclude that the identification of algal patterns and their driving forces can provide valuable information to aid bioassessment at the river system scale.

## 1. Introduction

Spatio-temporal patterns of lotic assemblages change substantially along anthropogenic stressor gradients (e.g., water quality pollution, hydrological alteration, and sediment erosion). Hence, understanding biological patterns and their driving forces can aid in the identification of the stressors affecting them [1]. For a given river system, the assessment of changes in lotic assemblage structures/patterns and functions responding to stressors, namely bioassessments, could not only indicate the causes of degradation in the biological condition, but also guide proper watershed management practices [2–4]. Therefore, in recent decades, lotic bioassessments have been increasingly included as foundational components of watershed management plans [5,6].

However, performing lotic bioassessments at the river system scale pose many challenges [7], especially for relatively small rivers of hundreds of kilometers in length and thousands of $km^2$ in watershed size. When bioassessments are performed at large scales, e.g., at regional or national scales [8,9], although both natural environmental factors (e.g., temperature, precipitation, geology,

and soils) and anthropogenic stressors change significantly, stressors gradients are usually strong enough to drive changes in biological patterns. Thus, significant stressor-response relationships can be developed, and major stressors affecting lotic ecosystems can be effectively detected while accounting for variation in natural environmental factors [10,11]. In contrast, when bioassessments are performed at the river system scale, associations between biological patterns and stressors are difficult to identify, because small-scale spaces may share similar features with natural environmental factors, resulting in relatively homogeneous background water chemistry [12]. Conversely, physical habitat variables, such as elevation, slope, substrate, water depth, current velocity, and riparian cover, do change considerably along stream orders [13]. Therefore, the distributional patterns of lotic assemblages in a river system respond to both longitudinal variation in physical habitat conditions and anthropogenic stressor gradients in the watershed [14–17]. Naturally, longitudinal variation in biological patterns will obscure stress–response relationships if the stressor gradients are not strong enough [18]. Besides the spatial confounding effects, the effects of anthropogenic stressors and physical habitat variables on lotic assemblages may change temporally, inducing distinctly temporal stress–response relationships [7,19].

Therefore, traditional bioassessment methods may lead to inaccurate results when they are used in river systems without additional analysis. For instance, biological indices that were initially developed to indicate ecological conditions, such as trophic or water quality status, would also be influenced by variation in the natural environment [20,21]. As for reference condition-based methods, assessments may be hard to perform within river system due to limited number of reference sites [22,23]. Reference sites may be relatively easy to find in upstream areas but are scarce or totally absent in downriver areas where there is extensive human activity. It is evident that upstream reference sites cannot represent downriver reference conditions. Under extreme conditions, human disturbances in upstream areas may also be so severe that only a few reference sites can be identified. Under such conditions, assessment accuracy becomes a knotty issue [24]. The application of alternative methods to explore associations between biological patterns and environmental variables could provide valuable information that aids in bioassessments at the river system scale; however, such attempts have been infrequent (but see [18]).

Lotic benthic algae, especially benthic diatoms, respond rapidly to environmental changes because of their nutrient requirements and their position at the base of lotic food webs. Hence, they can provide unique early warning signals of deteriorating conditions and have been extensively used in lotic assessments [25]. In the present study, we used a bottom-up strategy to study the associations between the spatio-temporal patterns of benthic diatom assemblages and water quality in a mountainous Chinese river system. We first performed hierarchical clustering to identify the spatial and temporal patterns of benthic diatom assemblages within the river system. Then, the random forest method was used to identify the contributions of the physical habitat and chemical variables to diatom patterns. Spatial differences in environmental variables were detected with nonparametric tests, and indicator taxa for diatom patterns were identified with an indicator species analysis. Meanwhile, several European diatom indices that are generally used to indicate lotic ecological conditions were also calculated. Linear models were fitted to identify potential driving forces for variation in diatom indices. It was assumed that (1) the contributions of the physical habitat and water quality variables to algal patterns in the river system could be distinguished by combining applications of several analysis methods. Therefore, these analyses may provide valuable aid to bioassessments, (2) European indices would lead to inaccurate assessments if they are used without modification. Our study demonstrates that more consideration of bioassessments within a river system is needed.

## 2. Materials and Methods

### 2.1. Study Area

The Chishui river (104°45′–106°51′ E, 27°20′–28°50′ N) is located in southwestern China [26,27]. The river originates from Zhenxiong County in northern Yunnan Province and runs through 13 counties/cities in Yunnan, Guizhou, and Sichuan province before flowing into the Yangtze River in Hejiang County. The main stream is 436.5 km long with a drainage area of 20,440 km$^2$. The elevation of the watershed ranges from 200 to 1800 m, with the terrain gently sloping from the southeast towards the northwest. This area has a subtropical continental monsoon climate with hot, wet summers and cool, dry winters. The annual average air temperature varies from 11.3 °C in the headwaters to 18.1 °C in the downstream sub-basins [28]. The annual precipitation is 1027.2 mm, and approximately 60% of rainfall occurs between June and September.

Although the Chishui river is the only tributary of the upper Yangtze River, where no dam has been built on the main stream, other types of human disturbance are popular [29]. Specifically, the upper reaches of the river suffer from the effects of domestic sewage and agricultural pollution. Villages are situated near the riverbank, and crops have been planted within the riparian zones because there is limited flatland in this mountainous area. The main human disturbances in the middle reach include wine companies, coal mining enterprises, and towns. In the downriver reaches, with the increased urbanization and industrialization, urban and industrial wastewater, combined with agricultural pollution, has had a substantial influence on water quality.

### 2.2. Benthic Diatom Sampling and Identification and Diatom Index Calculation

To account for the considerable temporal variation in assemblage compositions, we sampled benthic algae from 40 main stream river sites in April (the base flow period) and September (the high flow period) 2016 (Figure 1). During each sampling occasion, 15 samples were collected from the dominant substrate along a reach that was several times longer than the width of the sampling transects (approximately 50–200 m) at each site [30,31]. Generally, algae were sampled from randomly selected rocks (diameter range: 15–60 cm) with the sampling area confined using a circular lid (radius: 2.7 cm). For each rock, the surface within the lid was vigorously scrubbed using a nylon brush and rinsed three to four times with distilled water. Samples from the same site were combined into one composite sample, and the volume was recorded. For several downstream, non-wadable sites where the dominant substrate was sediment, a 60-mL syringe was used to collect the top 1 cm of soft sediment from the lid-delimited area. Sediment from the same site was combined. For wadable sites, algal samples were collected across the stream transects. In contrast, algae were sampled from shorelines at a water depth of <0.5 m at the non-wadable sites because algal growth is limited by deeper water with a higher turbidity [18]. All samples were preserved with 4% formalin in the field for further identification and enumeration.

In the laboratory, sediment samples were first rinsed with distilled water several times to separate algae from mud. After this pretreatment, diatoms were cleaned with concentrated sulfuric acid and concentrated nitric acid and mounted on a microscope slide with neutral balsam (Hushi Brand, Sinopharm Chemical Reagent Co., Ltd., Shanghai, China, refractive index: 1.52) [32]. Diatoms were enumerated and identified using a compound microscope (Olympus CX21: Olympus Optical Co., Tokyo, Japan) at 1000X. At least 600 valves were identified (most to the species level) following the taxonomic references [33–38]. Basionyms were examined according to the Catalogue of Diatom Names (Online Version, updated 19 September 2011, http://researcharchive.calacademy.org/research/diatoms/names/index.asp), and only the currently used names were retained. The number of each taxon presenting at each site was counted to determine the relative abundances. Algal samples and diatom slides were deposited in the lab for future reference.

Four European diatom indices that are widely used in lotic assessments were calculated with OMNIDIA 7 software V 4.2 [39]: the Sladecek index (SLA), the Specific Pollution sensitivity

Index (IPS: Indice de Polluo-sensibilité Spécifi que), the Biological Diatom Index (IBD: Indice Biologique Diatomées), and the Trophic Diatom Index (TDI). They were calculated based on the Zelinka–Marvan weighted average equation, in which the relative abundance of each diatom taxa in the community is combined with their ecological valency (sensitivity or optima) and indicator weights (tolerance) [40]. The indicator values (i.e., the sensitivity/optima values or the weights/tolerance values) for each taxon were determined a priori. The four indices differed in several respects, including the number of taxa used to develop the indices, the taxonomic resolution, the indicator values attributed to each taxon, and the information provided. Specifically, SLA is a saprobic index that measures the degree of organic enrichment in waters; IPS evaluates the general water quality by integrating organic pollution, salinity, and eutrophication; IBD was developed to indicate how much the level of organic and global pollution in a water body; and TDI is commonly used to assess the trophic status of rivers and streams [20,40].
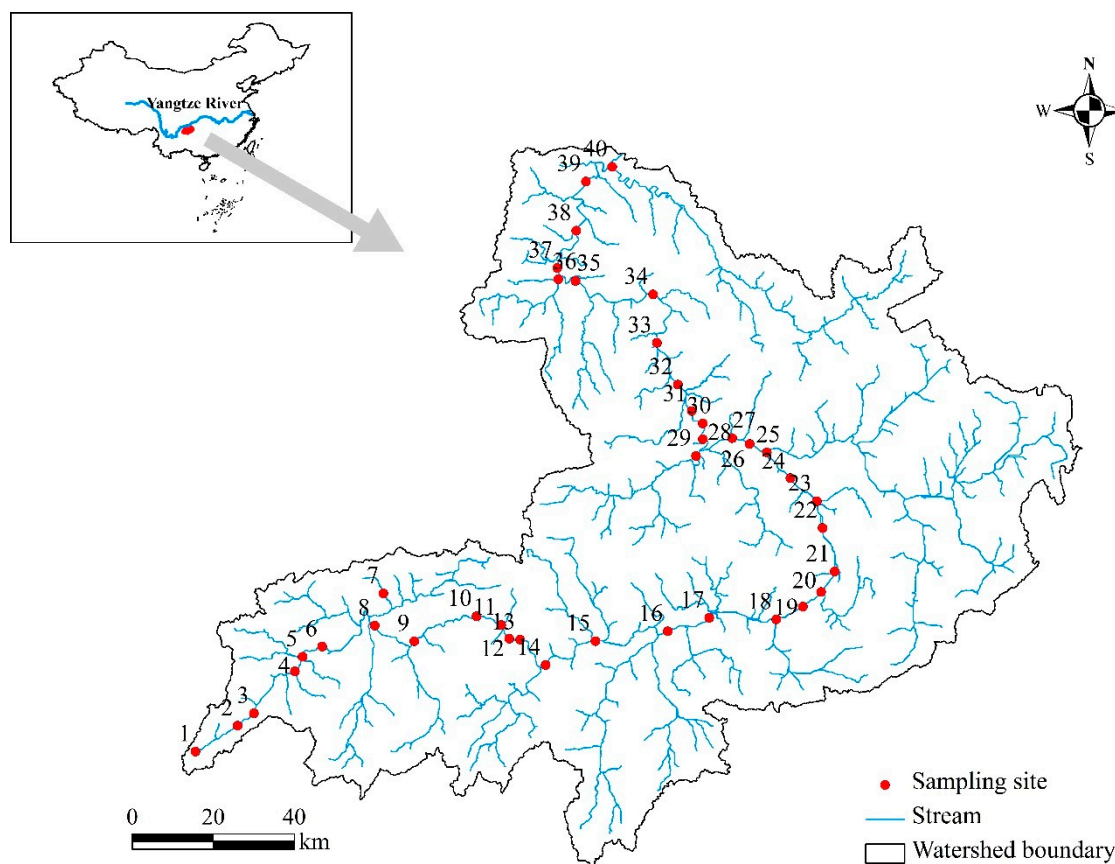


**Figure 1.** Locations of the study sites in the Chishui river network. Sites are numbered from the headwaters (1) to the river's mouth (40) along the flow direction.

*2.3. Environmental Variables*

For each site, the location (latitude, longitude, and elevation) and slope of the river reach were recorded with a GPS system and a gradiometer, respectively. The current velocity was measured with a current flow meter (FP211, Global Water). pH and conductivity were measured with a portable Yellow Springs Instrument (YSI) meter (Model 33, YSI, Incorporated, Yellow Springs, Greene County, OH, USA). A 600 mL stream water sample was collected and preserved in acidic conditions until chemical analysis. In the laboratory, concentrations of Ammoniacal nitrogen ($NH_4$–N) total nitrogen (TN), Total phosphorus (TP), and Chemical oxygen demand (COD) were measured following the standard methods recommended by the national water monitoring protocol [41].

### 2.4. Statistical Analysis

We first clustered sites with the species composition of diatom assemblages using agglomerative hierarchical clustering based on the Bray–Curtis dissimilarity with a flexible linkage of $\beta = -0.25$. Bray–Curtis dissimilarity conserves the object space and is commonly used in cluster preprocessing, and a flexible linkage of $\beta = -0.25$ has shown good performance in clustering [42]. Prior to clustering, taxa appearing in $\leq 5\%$ or $\geq 95\%$ of the sites were excluded because they contributed little to the classification [43], and relative abundance data were arcsine–square root transformed to down weight the influence of the most abundant taxa.

After the optimum number of clusters had been determined, the random forest method was conducted to develop a discrimination model for site classification. Random forest is a boosted tree method that prevents overfitting in classification and regression trees, and releases the strict requirements on data independence and normality of traditional linear methods [44]. This method has been proven to perform better than linear discriminant analysis and generalized dissimilarity modeling in classification analysis [45]. The diatom group for each site that was assigned by hierarchical clustering was used as the response variable, and nine physical and chemical variables that are important to algal distributional patterns were used as predictive variables: COD, $NH_4$-N, TN, TP, pH, conductivity, elevation, slope, and velocity. The classification performance was evaluated by the out-of-bag (OOB) error rate. The model was run 100 times, and the average error rate was calculated to offset the randomness in data selection. We successively removed each predictor that made the least contribution from the model, and at each step, the OOB error rate of the new model was compared to that of the model with previous predictors. The process stopped when the new model had a higher error rate than the old one, at which point the remaining predictors were considered to be the most important contributors to the site classification [44]. In addition, Kruskal–Wallis tests were performed to examine whether there were significant differences in these predictors among the site groups. The Mann–Whitney U test was adopted as a post-hoc test for pairwise comparison between site groups.

We further used the indicator species analysis (ISA) to identify which diatom taxa were important to individual site groups. ISA calculates an indicator value (IndVal) for each taxon by combing the relative abundances and frequencies within each classification group [46]. The indicator taxa for each site group are thereafter selected based on their specificity (the probability that a taxon occurs in the target site group) and fidelity (the probability of finding the species in sites belonging to the target site group) to the group. The IndVal ranges from 0 to 1, with 1 representing a perfect indication. The significance of IndVal was tested with the Monte Carlo randomization technique. Taxa with IndVal $\geq 0.5$ and $p < 0.05$ in permutation tests (1000 times) were considered to be indicators for their respective groups.

A stepwise multiple linear regression analysis was performed to identify the subset of environmental variables that best explained the observed variation in diatom indices. For the initial models, all nine physical and chemical variables were included as predictors, with one of the diatom indices as the response variable. The Akaike Information Criterion (AIC) was adopted to evaluate the improvement of the model when adding or dropping a predictor. After the final model had been fitted, quantile-quantile plots and plots of residuals versus fitted values were checked to determine whether the assumptions of linear regression had been fulfilled. Predictors with extreme values were transformed and the collinearity between predictors was checked before analysis. Regression models were fitted separately for the data from April and September.

All statistical analyses were undertaken in R 3.3.3 (R Core Team 2016). Agglomerative hierarchical clustering, random forest, and indicator species analysis were performed with the "Cluster", "randomForest", and "indicspecies" packages, respectively. Stepwise multiple linear regression was implemented with the "lm" and "step" functions from the base statistical package "stats".

## 3. Results

### 3.1. Site Classification

A total of 183 diatom taxa were observed in April (120) and September (107) of 2016. The predominant species, *Achnanthidium minutissimum* (Kützing) Czarnecki and *Cocconeis placentula* Ehrenberg, occurred at all sites and in both months, with average relative abundances of 31.9% and 20.3% in April and 44.3% and 16.7% in September, respectively. A further 11 taxa had an average relative abundance >1%: *Brebissonia lanceolata* (Agardh) Mahoney & Reimer, *Gomphonema parvulum* Kützing, *Luticola mutica* (Kützing) Mann, *Melosira varians* Agardh, *Navicula antonii* Lange-Bertalot & Rumrich, *Navicula cryptocephala* Kützing, *Navicula reichardtiana* Lange-Bertalot, *Nitzschia dissipata* (Kützing) Rabenhorst, *Nitzschia palea* (Kützing) Smith, *Planothidium lanceolatum* (Brébisson) Round et Bukhtiyarova, and *Rossithidium linearis* (Smith) Round et Bukhtiyarova. The average number of taxa that occurred in at study sites was 26 (with a range of 9–45) for April and 23 (9–42) for September. Generally, more taxa were observed at the middle and downriver sites than at the upstream sites.

The study sites were classified based on the community composition of diatom assemblages (Figure 2). The Davies-Bouldin clustering index, a commonly used measurement to determine the optimal number of the cluster, was minimized at 5 and 6. The 5-cluster structure was then chosen for its conciseness. The agglomerative coefficient of the 5-cluster structure, which describes the strength of the clustering structure obtained by a group's average linkage, was 0.82, indicating a reasonable clustering. The clustering results showed clear temporal variation (Figure 2). All April sites were assigned to group 1 (G1:19 sites) or group 2 (G2: 21 sites). G1 was mainly composed of upstream and middle-river sites, while G2 was composed of most downriver sites. September sites were clustered to the other three groups (G3–G5). Most upstream sites in September were assigned to G5 (14 sites), with middle-river sites to G3 (18 sites) and downriver sites to G4 (8 sites). Sites allocated to G1 in April were mainly assigned to G3 (36.8%) and G5 (57.9%) in September. In contrast, G2 sites were assigned to G3 (47.6%) and G4 (38.1%) in September.
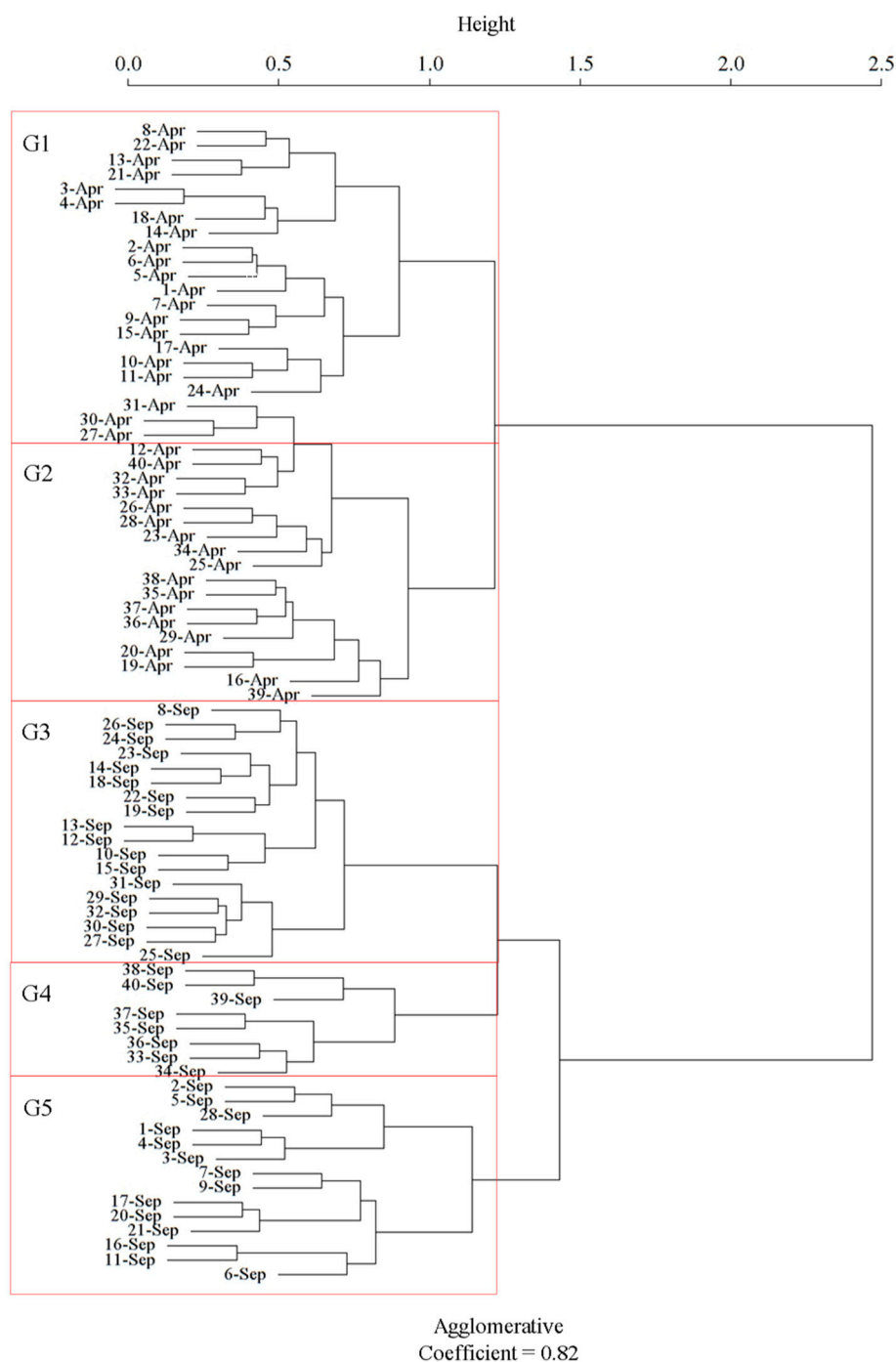
**Figure 2.** Agglomerative hierarchical clustering dendrogram for site classification. Sites are numbered from the headwaters (1) to the river's mouth (40) along the flow direction and labeled with sampling months. Apr: April, Sep: September. G1 to G5 represent the individual site groups.

### 3.2. Differences in Environmental Variables among the Site Groups

All environmental variables, except slope and velocity, showed significant differences among site groups (Kruskal–Wallis test, $p < 0.01$, Figure 3). Significant intergroup differences and temporal differences were revealed by the Mann–Whitney U tests. The COD concentrations in the April groups (G1 and G2 medians: 1.7 and 3.2 mg/dm$^3$, respectively) were lower than those in the September groups (medians for G3–G5: 9.59, 9.04, and 7.31 mg/dm$^3$, respectively). Besides this, the COD concentrations in G1 were lower than those in G2. The elevations of the G1 (median: 735 m) and G5

(median: 999.6 m) sites were higher than those of the other three groups of sites. Although G2, G3, and G4 all had relatively low elevations, they differed from one other. The G4 sites had the lowest elevations (median: 224.2 m), with higher elevations for the G2 sites (median: 288.4 m), and even higher for the G3 sites (median: 381.8 m). The TN concentrations in G1, G3, and G5 (medians: 3.81, 3.99, and 3.95 mg/dm$^3$, respectively) were higher than those in G2 and G4 (medians: 3.21 and 3.17 mg/dm$^3$, respectively). Conductivity, $NH_4$–N, and TP were higher in April than in September. The pH displayed opposite temporal changes.
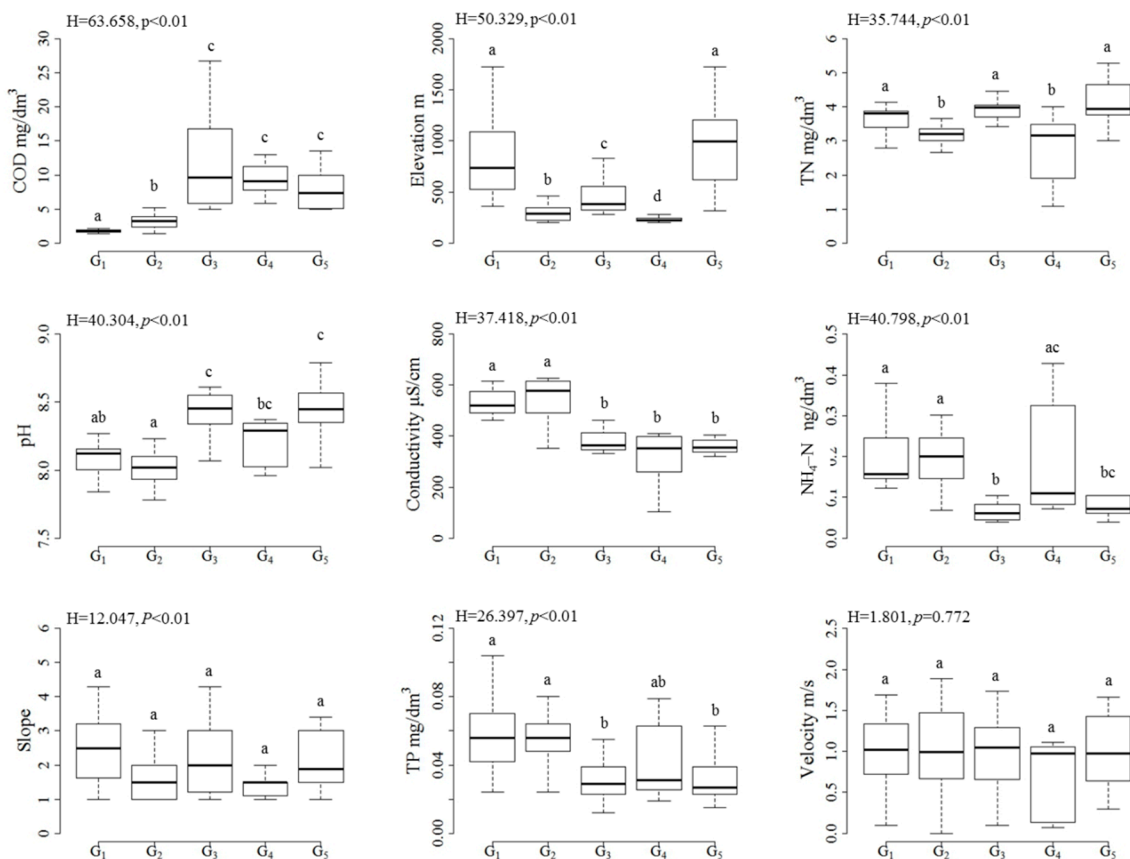


**Figure 3.** Boxplots displaying the differences in the environmental variables among the five site groups. G1 to G5 under the x-axis represent the site groups identified by cluster analysis. The boxes, horizontal lines in the boxes, and whiskers represent the interquartile range, median, and value range of the variables, respectively. The H value and the *p*-value are the results of Kruskal–Wallis tests among the site groups. The lower-case letters above whiskers are the results of pairwise comparisons, with the same letters indicating no difference at $p \leq 0.05$.

The full random forest model that included all nine environmental variables and the sub-model with COD, elevation, and TN had similar OOB estimates of classification error rate (21.36% for the full model versus 20.94% for the sub-model). COD was the most important determinant of site classification (Figure 4), followed by elevation and TN.
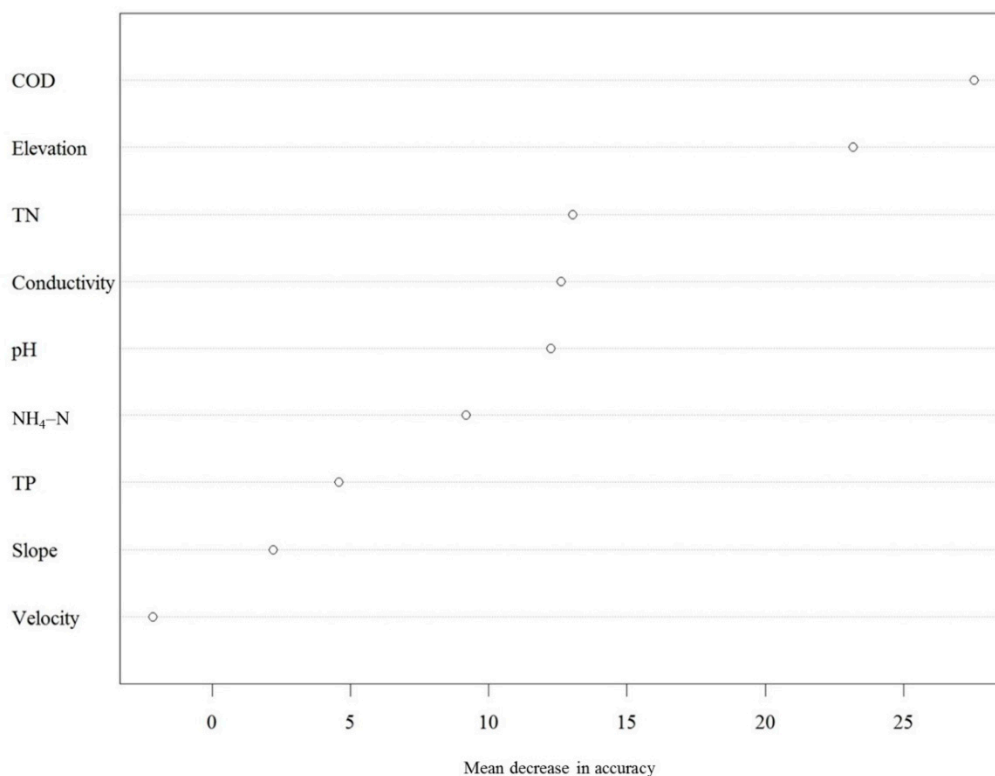
**Figure 4.** Contributions of physical and chemical variables to the site classification in the random forest analysis. The mean decrease in accuracy represents the importance of each variable to the classification. COD: Chemical oxygen demand, TN: total nitrogen, $NH_4$–N: Ammoniacal nitrogen, TP: Total phosphorus.

### 3.3. Indicator Taxa for Site Groups

In total, 49 indicator taxa were identified for the five site groups (Table 1). There were two indicator taxa from different genera for G1. G2 had the highest number of indicator taxa (22) from 16 genera. Sixteen indicator taxa were identified for G3, with most taxa from the *Navicula* sensu lato (7) and *Cymbella* (4) genera. G4 had eight indicator taxa, five from the motile genera *Nitzschia* and *Navicula* [47]. *Cocconeis placentula* was the only indicator for G5. Overall, approximately 70% of all indicator taxa came from five genera: *Navicula* (10), *Nitzschia* (6), *Cymbella* (8), *Achnanthes* sensu lato (4), and *Amphora* sensu lato (5). Taxa from the motile genera *Nitzschia* or *Navicula* mainly occurred in G3 and G4.

**Table 1.** Indicator values (IndVal) of selected diatom taxa for individual site groups. Site groups were identified by a cluster analysis. The numbers in brackets following site group names are the numbers of indicator taxa. A refers to species specificity to a site group, which is the probability that a taxon occurs in a particular target site group; B refers to species fidelity, which is the probability of finding the species at the sites belonging to the target site group.

| Taxa | A | B | lndVal | *p*-Value |
|---|---|---|---|---|
| **G1 (N = 2)** | | | | |
| *Rossithidium linearis* (Smith) Round et Bukhtiyarova | 0.743 | 1.000 | 0.862 | 0.005 |
| *Nitzschia acicularis* (Kützing) Smith | 0.645 | 0.421 | 0.521 | 0.020 |
| **G2 (N = 22)** | | | | |
| *Encyonema muelleri* (Hustedt) Mann | 0.970 | 0.571 | 0.744 | 0.005 |
| *Melosira granulata* (Ehrenberg) Ralfs | 0.966 | 0.571 | 0.743 | 0.005 |
| *Ulnaria ulna* (Nitzsch) Compère | 0.635 | 0.857 | 0.738 | 0.005 |
| *Nitzschia pseudofonticola* Hustedt | 0.757 | 0.714 | 0.735 | 0.005 |
| *Gomphonema clavatum* Agardh | 0.806 | 0.667 | 0.733 | 0.005 |
| *Diatoma vulgare* Bory | 0.730 | 0.714 | 0.722 | 0.010 |
| *Halamphora coffeaeformis* (Agardh) Levkov | 0.890 | 0.524 | 0.683 | 0.005 |
| *Navicula cryptocephala* Kützing | 0.478 | 0.952 | 0.674 | 0.005 |
| *Cymbella tumidula* Grunow | 0.945 | 0.476 | 0.671 | 0.005 |
| *Melosira varians* Agardh | 0.435 | 1.000 | 0.659 | 0.010 |
| *Achnanthes rossii* Hustedt | 0.977 | 0.429 | 0.647 | 0.005 |
| *Diatoma moniliformis* (Kützing) Williams | 0.786 | 0.524 | 0.642 | 0.005 |
| *Amphora pediculus* (Kützing) Grunow | 0.706 | 0.571 | 0.635 | 0.005 |
| *Encyonopsis behrei* (Foged) Krammer & Metzeltin | 0.719 | 0.524 | 0.614 | 0.010 |
| *Cocconeis pediculus* Ehrenberg | 0.395 | 0.952 | 0.613 | 0.010 |
| *Delicata gadjiana* (Maillard) Krammer | 0.943 | 0.381 | 0.599 | 0.005 |
| *Cyclotella meneghiniana* Kützing | 0.894 | 0.381 | 0.583 | 0.020 |
| *Encyonema silesiacum* (Bleisch) Mann | 0.884 | 0.381 | 0.580 | 0.010 |
| *Fragilaria vaucheriae* (Grunow) Jørgensen | 1.000 | 0.333 | 0.577 | 0.010 |
| *Cyclotella* sp. | 1.000 | 0.286 | 0.535 | 0.010 |
| *Gomphonema olivaceum* (Hornemann) Ehrenberg | 0.594 | 0.476 | 0.532 | 0.010 |
| *Achnanthes impexiformis* Lange-Bertalot | 0.892 | 0.286 | 0.505 | 0.015 |
| **G3 (N = 16)** | | | | |
| *Luticola mutica* (Kützing) Mann | 0.929 | 0.944 | 0.937 | 0.005 |
| *Cymbella tumida* (Brébisson ex Kützing) Van Heurck | 0.913 | 0.944 | 0.929 | 0.005 |
| *Cymbella turgidula* Grunow | 0.862 | 0.889 | 0.875 | 0.005 |
| *Cymbella peraffinis* Tynni | 0.866 | 0.778 | 0.821 | 0.005 |
| *Encyonema lacustre* (Agardh) Mann | 0.804 | 0.778 | 0.791 | 0.005 |
| *Cymbella saxicola* Bíly & Marvan | 1.000 | 0.611 | 0.782 | 0.005 |
| *Navicula antonii* Lange-Bertalot & Rumrich | 0.551 | 1.000 | 0.742 | 0.005 |
| *Luticola stigma* (Patrick) Johansen | 0.990 | 0.556 | 0.742 | 0.005 |
| *Nitzschia dissipata* (Kützing) Rabenhorst | 0.538 | 1.000 | 0.733 | 0.005 |
| *Amphiprora paludosa* Smith | 0.942 | 0.500 | 0.686 | 0.005 |
| *Gomphonema subclavatum* (Grunow) Grunow | 0.533 | 0.722 | 0.621 | 0.015 |
| *Gyrosigma scalproides* (Rabenhorst) Cleve | 0.510 | 0.722 | 0.607 | 0.030 |
| *Amphora bullatoides* Hohn & Hellerman | 0.702 | 0.500 | 0.592 | 0.020 |
| *Navicula cryptotenella* Lange-Bertalot | 0.560 | 0.556 | 0.558 | 0.020 |
| *Navicula cryptofallax* Lange-Bertalot & Hofmann | 0.632 | 0.444 | 0.530 | 0.015 |
| *Navicula rostellata* Kützing | 0.903 | 0.278 | 0.501 | 0.010 |
| **G4 (N = 8)** | | | | |
| *Planothidium lanceolatum* (Brébisson) Lange-Bertalot | 0.658 | 1.000 | 0.811 | 0.005 |
| *Halamphora montana* (Krasske) Levkov | 0.910 | 0.625 | 0.754 | 0.005 |
| *Navicula reichardtiana* Lange-Bertalot | 0.638 | 0.875 | 0.747 | 0.005 |
| *Nitzschia palea* (Kützing) Smith | 0.624 | 0.875 | 0.739 | 0.010 |
| *Navicula symmetrica* Patrick | 0.658 | 0.750 | 0.703 | 0.005 |
| *Amphora cymbamphora* Cholnoky | 0.708 | 0.625 | 0.665 | 0.005 |
| *Nitzschia valdecostata* Lange-Bertalot & Simonsen | 0.822 | 0.375 | 0.555 | 0.005 |
| *Nitzschia minuta* Bleisch | 0.606 | 0.500 | 0.550 | 0.020 |
| **G5 (N = 1)** | | | | |
| *Cocconeis placentula* Ehrenberg | 0.348 | 1.000 | 0.590 | 0.010 |

*3.4. Environmental Variables Best Explaining the Variation in European Diatom Indices*

More than 80% of observed taxa were used to calculate the TDI (April: 81.5%; September: 84.3%) and IPS (April: 80.6%; September: 83.5%). By comparison, approximately 65% and 50% of taxa were included for calculating the IBD (April: 63.4%; September: 64.9%) and the SLA (April: 51.0%; September: 52.9%), respectively. The stepwise multiple linear regression analysis revealed distinct relationships between the four European diatom indices and environmental variables (Table 2). In April, approximately 20% of the variation in SLA was explained by $NH_4$–N and current velocity. In comparison, TP, pH, elevation, and current velocity explained more than 40% of the variation in this index in September. Elevation was the only important environmental variable for the TDI in both months, with a relatively stronger relationship in April. Variation in the IBD was also explained only by elevation in April. Variation in the IPS in both months and the IBD in September was not associated with any environmental variable.

**Table 2.** Environmental variables that best explain the variation in the European diatom indices revealed by the stepwise multiple linear regression analysis. SLA: the Sladecek index, IBD: the Biological Diatom Index, TDI: the Trophic Diatom Index. Coeff.: regression coefficient, $R^2_{adj}$: Adjusted R-squared.

| Month | Diatom Indices | Predictors | Coeff. | *p*-Value | *F* Statistic | *p*-Value | $R^2_{adj}$ |
|---|---|---|---|---|---|---|---|
| April | SLA | Intercept | 11.422 | <0.0001 | 5.624 | 0.0074 | 0.192 |
| | | $NH_4$–N * | −5.793 | 0.0071 | | | |
| | | Velocity * | 1.578 | 0.0268 | | | |
| | IBD | Intercept | 4.219 | 0.0002 | 9.875 | 0.0032 | 0.185 |
| | | Elevation | 0.004 | 0.0032 | | | |
| | TDI | Intercept | 58.140 | <0.0001 | 7.029 | 0.0116 | 0.134 |
| | | Elevation | −0.012 | 0.0116 | | | |
| September | SLA | Intercept | 23.440 | 0.0002 | 8.325 | <0.0001 | 0.429 |
| | | TP * | −27.330 | <0.0001 | | | |
| | | pH | −1.438 | 0.0353 | | | |
| | | Elevation | −0.001 | 0.0215 | | | |
| | | Velocity * | 3.852 | 0.0068 | | | |
| | TDI | Intercept | 54.574 | <0.0001 | 4.107 | 0.0498 | 0.074 |
| | | Elevation | −0.008 | 0.0498 | | | |

*: variable with $\log_{10}(x + 1)$ transformation.

## 4. Discussion

*4.1. Associations between Algal Patterns and Water Quality Gradients*

We found that benthic diatom assemblages in the Chishui river system could be classified into two or three groups depending on the sampling month. In some cases, sites from different longitudinal locations were assigned to the same group. For instance, both upstream and downriver sites belonged to group 1, and group 3 was composed of sites from up-, middle-, and downriver sites. This implies that diatom patterns in this river system are not solely determined by the longitudinal location as observed in some other rivers (Figure 2) [18]. We further found that COD, elevation, and TN were the most important environmental variables affecting the spatio-temporal patterns of diatom assemblages (Figure 3), indicating that both natural habitat and water quality variables are important to diatom patterns in this river system.

The COD was mainly produced by the hundreds of wine enterprises that are located in the middle reach of the river system [48]. These enterprises are distributed near the river for convenience of extracting water. The river's water quality has been severely degraded by wastewater emissions from these enterprises under the current management standards [49]. Organic pollutants that are discharged into rivers can, through oxidation, be converted into nutrients that support algal growth. Other studies have also found that COD plays an important role in diatom patterns in highly

disturbed areas [50,51]. Similar to most other studies [52,53], TN was shown to have important effects on diatom patterns in the Chishui river. TN in this river was mainly generated from domestic and agricultural pollution. Contrary to COD, TN concentrations were higher at the upstream sites. This finding is in accordance with the fact that stream banks in upstream reaches of this mountainous watershed are generally occupied by buildings and farmlands which induce severe domestic and agricultural pollution. Elevation, an important natural variable that shapes biological patterns globally [54], also had important effects on diatom patterns in the Chishui river. This variable is a proxy factor for temperature, energy, precipitation, and other variables that are essential to algal distribution. In the present study, elevation decreased longitudinally from the headwaters to the river's mouth. Hence, diatom assemblages also changed significantly between high- and low-elevation sites, irrespective of the sampling month (G1 versus G2 in April and G5 versus G4 in September).

Diatom assemblages in April and September were separated by the first level of clustering (Figure 2), indicating that temporal changes in diatom patterns were more significant than those in spatial variation. The results of the nonparametric tests imply that temporal changes in diatom patterns were mainly associated with changes in the COD concentration. COD concentrations in September were significantly higher than that in April, which could be attributed to increased terrestrial inputs of organic pollutants through rainfall events.

Indicator taxa provided further evidence for associations between diatom patterns and water quality. *Nitzschia acicularis*, a eutrophic and *α*-mesosaprobous species [55], was indicative to G1, implying a potential water quality problem in this group. Taxa indicating the eutrophic condition, including *Amphora pediculus*, *Cocconeis pediculus*, *Cyclotella meneghiniana*, *Cymbella tumida*, *Cymbella turgidula*, *Diatoma moniliformis*, *Encyonema lacustre*, *Gomphonema clavatum*, *Gyrosigma scalproides*, *Halamphora coffeaeformis*, *Halamphora montana*, *Luticola mutica*, *Luticola stigma*, *Melosira varians*, *Navicula antonii*, *Navicula symmetrica*, *Nitzschia dissipata*, *Nitzschia palea*, *Nitzschia pseudofonticola*, and *Planothidium lanceolatum*, were more representative of the low-elevation groups (G2, G3 and G4). The autecology of most indicator taxa suggests that water quality degradation is extensive in this watershed, and the diatoms were an effective indicator of such type of disturbance. Besides the water quality, the indicator taxa also revealed that there is likely a sediment gradient in the Chishui river system that has increased the propitiation of motile genus *Navicula* and *Nitzschia* in low-elevation groups 3 and 4. Smucker et al. [56] demonstrated that an increase in these motile taxa also indicates an increase in human disturbance represented as the watershed's impervious cover percentage.

## 4.2. Potential Applications

Our analysis is helpful for a more accurate understanding of ecological conditions the Chishui river system. We found that the variation in European diatom indices was only partly associated (for the SLA), or unassociated (for the IPS, IBD, and TDI), with water quality variables. Other researchers have obtained similar results and argued that diatom indices derived from the region under study would perform better than that of foreign indices [20,57]. We did not develop diatom indices for the Chishui river because only the main river stream was surveyed. Nevertheless, it would be worthwhile to attempt this if funding for a more extensive survey becomes available in the future. Using a reference condition-based assessment, Chi et al. [29] developed a macroinvertebrate-based multi-metric index for the Chishui river. Compared with other studies [8,58], they set relatively low-quality criteria (TP < 0.20 mg/L, $NH_4$–N < 1 mg/L, and COD < 6 mg/L) for screening reference sites. They found that <10% of sites were in "excellent" condition, with some headwater and upstream sites in "fair" or "poor" condition. Their results coincide with our finding that upstream sites were also polluted. As pollution was extensive in the watershed, we did not adopt a reference condition-based bioassessment. Alternatively, we classified sites directly with community compositions of benthic diatoms and identified the environmental variables driving algal patterns. Although we could not measure specific ecological conditions for individual sites with this approach, we did find that an environmental stressor, water quality degradation, affected the river ecosystem.

Thus, the combination of our findings with those of Chi et al. [29] allows a more comprehensive and accurate understanding of this river system's ecological conditions.

Diatom patterns, together with their associations with environmental variables in the Chishui river system may be used to guide management practices. First, the upper reaches of this river ecosystem were mainly influenced by nitrogen pollution generated by domestic and agricultural activities on the river banks. Riparian vegetation zones are important buffering and filtering zones for terrestrial inputs of nutrients [59]; therefore, they urgently need to be recovered to protect the water quality in this area. As for the middle- and downriver reaches where the COD loadings were high due to the dense distribution of wine enterprises, more strict management might be considered. For instance, to protect water quality to sustain the high quality of the wine produced here, both the number of wine enterprises and the wastewater emission standards could be revised [48]. Secondly, the COD and TN loadings were shown to be higher in September than in April. This implies that more terrestrial contaminants were transferred to the river through rainfall. Therefore, more effective wastewater collection systems could be developed or updated for pollution control during the rainy season.

In conclusion, both physical habitat variable of elevation and water quality variables of COD and TN were important to the spatio-temporal patterns of diatom assemblages in the Chishui river system. Combining the application of hierarchical clustering, discriminant analysis, indicator species analysis, and nonparametric tests was effective in detecting the associations between algal patterns and water quality. Conversely, the use of European diatom indices would lead to a biased assessment. Overall, our study demonstrates that exploring associations between lotic biological patterns and environmental variables can provide valuable information to aid in bioassessments, especially for river systems where there is substantial natural variation in habitat conditions or where there has been extensive human disturbance.

**Author Contributions:** T.T. and Z.X. conceived and designed the research; S.G., T.T., and Z.X. performed the fieldwork; S.G. conducted the taxonomic identification and enumeration of diatoms; S.M. analyzed the data and wrote the manuscript; T.T., H.D., and Z.X. contributed to the discussion; and all authors reviewed the manuscript before submission.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yuan, L.; Norton, S.B. Assessing the relative severity of stressors at a watershed scale. *Environ. Monit. Assess.* **2004**, *98*, 323–349. [CrossRef]
2. Flotemersch, J.E.; Leibowitz, S.G.; Hill, R.A.; Stoddard, J.L.; Thoms, M.C.; Tharme, R.E. A watershed integrity definition and assessment approach to support strategic management of watersheds. *River Res. Appl.* **2016**, *32*, 1654–1671. [CrossRef]
3. Jähnig, S.; Tonkin, J.; Gies, M.; Domisch, S.; Hering, D.; Haase, P. Severity multipliers as a methodology to explore potential effects of climate change on stream bioassessment programs. *Water* **2017**, *9*, 188. [CrossRef]
4. Zhang, M.; Muñoz-Mas, R.; Martínez-Capel, F.; Qu, X.; Zhang, H.; Peng, W.; Liu, X. Determining the macroinvertebrate community indicators and relevant environmental predictors of the Hun-Tai River Basin (Northeast China): A study based on community patterning. *Sci. Total Environ.* **2018**, *634*, 749–759. [CrossRef]
5. Hering, D.; Borja, A.; Carstensen, J.; Carvalho, L.; Elliott, M.; Feld, C.K.; Heiskanen, A.-S.; Johnson, R.K.; Moe, J.; Pont, D.; et al. The European Water Framework Directive at the age of 10: A critical review of the achievements with recommendations for the future. *Sci. Total Environ.* **2010**, *408*, 4007–4019. [CrossRef]
6. Scrimgeour, G.J.; Hvenegaard, P.J.; Tchir, J. Cumulative industrial activity alters lotic fish assemblages in two boreal forest watersheds of Alberta, Canada. *Environ. Manag.* **2008**, *42*, 957–970. [CrossRef]

7. Mykrä, H.; Saarinen, T.; Tolkkinen, M.; McFarland, B.; Hämäläinen, H.; Martinmäki, K.; Kløve, B. Spatial and temporal variability of diatom and macroinvertebrate communities: How representative are ecological classifications within a river system? *Ecol. Indic.* **2012**, *18*, 208–217. [CrossRef]

8. Herlihy, A.T.; Paulsen, S.G.; Sickle, J.V.; Stoddard, J.L.; Hawkins, C.P.; Yuan, L.L. Striving for consistency in a national assessment: The challenges of applying a reference-condition approach at a continental scale. *J. N. Am. Benthol. Soc.* **2008**, *27*, 860–877. [CrossRef]

9. Hering, D.; Johnson, R.K.; Kramm, S.; Schmutz, S.; Szoszkiewicz, K.; Verdonschot, P.F.M. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: A comparative metric-based analysis of organism response to stress. *Freshwat. Biol.* **2006**, *51*, 1757–1785. [CrossRef]

10. Cao, Y.; Hawkins, C.P.; Olson, J.; Kosterman, M.A. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. *J. N. Am. Benthol. Soc.* **2007**, *26*, 566–585. [CrossRef]

11. Tang, T.; Stevenson, R.J.; Infante, D.M. Accounting for regional variation in both natural environment and human disturbance to improve performance of multimetric indices of lotic benthic diatoms. *Sci. Total Environ.* **2016**, *568*, 1124–1134. [CrossRef]

12. Olson, J.R.; Hawkins, C.P. Developing site-specific nutrient criteria from empirical models. *Freshw. Sci.* **2013**, *32*, 719–740. [CrossRef]

13. Vannote, R.L.; Minshall, G.W.; Cummins, K.W.; Sedell, J.R.; Cushing, C.E. The river continuum concept. *Can. J. Fish. Aquat. Sci.* **1980**, *37*, 130–137. [CrossRef]

14. Bellinger, B.J.; Angradi, T.R.; Bolgrien, D.W.; Jicha, T.M.; Hill, B.H.; Reavie, E.D. Longitudinal variation and response to anthropogenic stress in diatom assemblages of the Lower Mississippi River, USA. *River Syst.* **2013**, *21*, 29–54. [CrossRef]

15. Ferreira, F.; Petrere, M. The fish zonation of the Itanhaém river basin in the Atlantic Forest of southeast Brazil. *Hydrobiologia* **2009**, *636*, 11–34. [CrossRef]

16. Angradi, T.R.; Pearson, M.S.; Jicha, T.M.; Taylor, D.L.; Bolgrien, D.W.; Moffett, M.F.; Blocksom, K.A.; Hill, B.H. Using stressor gradients to determine reference expectations for great river fish assemblages. *Ecol. Indic.* **2009**, *9*, 748–764. [CrossRef]

17. Jiang, X.; Xiong, J.; Qiu, J.; Wu, J.; Wang, J.; Xie, Z. Structure of macroinvertebrate communities in relation to environmental variables in a Subtropical Asian river system. *Int. Rev. Hydrobiol.* **2010**, *95*, 42–57. [CrossRef]

18. Pan, Y.; Hughes, R.M.; Herlihy, A.T.; Kaufmann, P.R. Non-wadeable river bioassessment: Spatial variation of benthic diatom assemblages in Pacific Northwest rivers, USA. *Hydrobiologia* **2012**, *684*, 241–260. [CrossRef]

19. Wu, N.; Faber, C.; Sun, X.; Qu, Y.; Wang, C.; Ivetic, S.; Riis, T.; Ulrich, U.; Fohrer, N. Importance of sampling frequency when collecting diatoms. *Sci. Rep.* **2016**, *6*, 36950. [CrossRef]

20. Lavoie, I.; Hamilton, P.B.; Wang, Y.K.; Dillon, P. J.; Campeau, S. A. comparison of stream bioassessment in Québec (Canada) using six European and North American diatom-based indices. In *Diatom Taxonomy, Ultrastructure and Ecology: Modern Methods and Timeless Questions*; Kociolek, J.P., Theriot, E.C., Stevenson, R.J., Eds.; Schweizerbart Science Publishers: Stuttgart, Germany, 2009; pp. 37–56. ISBN 9783443510572.

21. Bere, T. Challenges of diatom-based biological monitoring and assessment of streams in developing countries. *Environ. Sci. Pollut. Res.* **2016**, *23*, 5477–5486. [CrossRef]

22. Kilgour, B.W.; Stanfield, L.W. Hindcasting reference conditions in streams. In *Landscape Influences on Stream Habitats and Biological Assemblages*; Hughes, R., Wang, L., Seelbach, P., Eds.; American Fisheries Society Symposium: Bethesda, MD, USA, 2006; pp. 623–639.

23. Birk, S.; van Kouwen, L.; Willby, N. Harmonising the bioassessment of large rivers in the absence of near-natural reference conditions—A case study of the Danube River. *Freshw. Biol.* **2012**, *57*, 1716–1732. [CrossRef]

24. Hawkins, C.P.; Olson, J.R.; Hill, R.A. The reference condition: Predicting benchmarks for ecological and water-quality assessments. *J. N. Am. Benthol. Soc.* **2010**, *29*, 312–343. [CrossRef]

25. Stevenson, R.; Pan, Y.; Van Dam, H. Assessing environmental conditions in rivers and streams with diatoms. In *The Diatoms: Applications for the Environmental and Earth Sciences*, 2nd ed.; Smol, J.P., Stoermer, E.F., Eds.; Cambridge University Press: Cambridge, UK, 2010; ISBN 9780521509961.

26. Huang, Z. On the development and conservation for the resources and ecology of Chishui river basin. *Res. Environ. Yangtze Basin* **2003**, *12*, 333–339. (In Chinese) [CrossRef]

27. Wu, J.; Zhao, H.; Miao, Z.; Chen, Y.; Zhang, F.; Wang, J. Status and conservation of fish resources in the Chishui River. *Biodivers. Sci.* **2010**, *18*, 162–168. (In Chinese)

28. Wang, Z.; Jiang, L.; Huang, M.; Zhang, C.; Yu, X. Biodiversity status and its conservation strategy in the Chishui river basin. *Res. Environ. Yangtze Basin* **2007**, *16*, 175–180. (In Chinese) [CrossRef]

29. Chi, S.; Gong, Y.; Wang, H.; Zheng, J.; Hu, J.; Hu, J.; Dong, F. A pilot macroinvertebrate-based multimetric index (MMI-CS) for assessing the ecological status of the Chishui River basin, China. *Ecol. Indic.* **2017**, *83*, 84–95. [CrossRef]

30. Tang, T.; Tang, T.; Tan, L.; Gu, Y.; Jiang, W.; Cai, Q. Identifying community thresholds for lotic benthic diatoms in response to human disturbance. *Sci. Rep.* **2017**, *7*, 4134. [CrossRef]

31. Tang, T.; Jia, X.; Jiang, W.; Cai, Q. Multi-scale temporal dynamics of epilithic algal assemblages: Evidence from a Chinese subtropical mountain river network. *Hydrobiologia* **2016**, *770*, 289–299. [CrossRef]

32. Hu, H.; Wei, Y. *The Freshwater Algae of China: Systematics, Taxonomy and Ecology*; Science Press: Beijing, China, 2006; ISBN 7030166337.

33. Qi, Y. *Flora Algarum Sinicarum Aquae Dulcis. Vol. 4. Bacillariophyta Centreae*; Science Press: Beijing, China, 1995; ISBN 7030044223.

34. Shi, Z. *Flora Algarum Sinicarum Aquae Dulcis, Tomus 12, Bacillariophyta-Ghophonemaceae*; Science Press: Beijing, China, 2004; ISBN 9787030133366. (In Chinese)

35. Krammer, K. *Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats. vol. 1: The Genus Pinnularia*; A.R.G.Gantner Verlag, K.G.: Ruggell, Liechtenstein, 2000; ISBN 978-3-904144-24-7.

36. Krammer, K. *Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats. Vol. 3: Cymbella*; A.R.G.Gantner Verlag, K.G.: Ruggell, Liechtenstein, 2002; ISBN 978-3-904144-84-1.

37. Krammer, K. *Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats. Vol. 4: Cymbopleura, Delicata, Navicymbula, Gomphocymbellopsis, Afrocymbella*; A.R.G. Gantner Verlag, K.G: Ruggell, Liechtenstein, 2003; ISBN 978-3-904144-99-5.

38. Lange-Bertalot, H.; Bak, M.; Witkowski, A.; Tagliaventi, N. *Diatoms of Europe: Diatoms of the European Inland Waters and Comparable Habitats. Vol. 2: Navicula Sensu Stricto, 10 Genera Separated from Navicula Sensu Lato, Frustulia*; A.R.G. Gantner Verlag, K.G.: Ruggell, Liechtenstein, 2001; ISBN 978-3-904144-78-0.

39. Lecointe, C.; Coste, M.; Prygiel, J. "Omnidia": Software for taxonomy, calculation of diatom indices and inventories management. *Hydrobiologia* **1993**, *269*, 509–513. [CrossRef]

40. Prygiel, J.; Coste, M.; Bukowska, J. Review of the major diatom-based techniques for the quality assessment of rivers-State of the art in Europe. In *Use of Algae for Monitoring Rivers III*; Prygiel, J., Whitton, B.A., Buckowska, J., Eds.; Agence de l'Eau Artois-Picar-die Press: Douai Cedex, France, 1999; ISBN 9782950208354.

41. Wei, F. *Monitoring and Analysis Methods of Water and Wastewater*; Chinese Environmental Sciences Publishment: Beijing, China, 2002.

42. Tornés, E.; Leira, M.; Sabater, S. Is the biological classification of benthic diatom communities concordant with ecotypes? *Hydrobiologia* **2012**, *695*, 43–55. [CrossRef]

43. Hawkins, C.P.; Vinson, M.R. Weak correspondence between landscape classifications and stream invertebrate assemblages: Implications for bioassessment. *J. N. Am. Benthol. Soc.* **2000**, *19*, 501–517. [CrossRef]

44. Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef]

45. Snelder, T.; Ortiz, J.; Booker, D.; Lamouroux, N.; Pella, H.; Shankar, U. Can bottom-up procedures improve the performance of stream classifications? *Aquat. Sci.* **2012**, *74*, 45–59. [CrossRef]

46. Dufrêne, M.; Legendre, P. Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecol. Monogr.* **1997**, *67*, 345–366. [CrossRef]

47. Stevenson, R.J.; Bahls, L.L. *Periphyton Protocols*; US Environmental Protection Agency: Washington, DC, USA, 1999; pp. 6–22.

48. Zhang, C.; Wang, Y.; Qiao, H.; Sun, L.; Dong, L. The impacts of liquor enterprises on water environment of Chishui River Basin. *Water Resour. Protect.* **2015**, *31*, 62–66. (In Chinese) [CrossRef]

49. Teng, Z.; Ding, A.; Li, Y.; Jing, R.; Liu, Y. Sources of water pollution and their spatio temporal variations in the upper reach of the Chishui River. *J. Beijng Norm. Univ. (Nat. Sci.)* **2016**, *52*, 3222–3327. (In Chinese) [CrossRef]

50. Soininen, J. Responses of epilithic diatom communities to environmental gradients in some Finnish rivers. *Int. Rev. Hydrobiol.* **2002**, *87*, 11–24. [CrossRef]

51.  Tan, X.; Xia, X.; Zhao, Q.; Zhang, Q. Temporal variations of benthic diatom community and its main influencing factors in a subtropical river, China. *Environ. Sci. Pollut. R.* **2014**, *21*, 434–444. [CrossRef]

52.  Soininen, J. Environmental and spatial control of freshwater diatoms—A review. *Diatom Res.* **2007**, *22*, 473–490. [CrossRef]

53.  Damanik-Ambarita, M.; Everaert, G.; Forio, M.; Nguyen, T.; Lock, K.; Musonge, P.; Suhareva, N.; Dominguez-Granda, L.; Bennetsen, E.; Boets, P.; et al. Generalized linear models to identify key hydromorphological and chemical variables determining the occurrence of macroinvertebrates in the Guayas River Basin (Ecuador). *Water* **2016**, *8*, 297. [CrossRef]

54.  Tonkin, J.D.; Shah, T.R.D.; Shah, D.N.; Hoppeler, F.; Jähnig, S.C.; Pauls, S.U. Metacommunity structuring in Himalayan streams over large elevational gradients: The role of dispersal routes and niche characteristics. *J. Biogeogr.* **2017**, *44*, 62–74. [CrossRef]

55.  Van Dam, H.; Mertens, A.; Sinkeldam, J. A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. *Neth. J. Aquat. Ecol.* **1994**, *28*, 117–133. [CrossRef]

56.  Smucker, N.J.; Detenbeck, N.E.; Morrison, A.C. Diatom responses to watershed development and potential moderating effects of near-stream forest and wetland cover. *Freshw. Sci.* **2013**, *32*, 230–249. [CrossRef]

57.  Tan, X.; Zhang, Q.; Burford, M.A.; Sheldon, F.; Bunn, S.E. Benthic diatom based indices for water quality assessment in two subtropical streams. *Front. Microbiol.* **2017**, *8*, 601. [CrossRef]

58.  Huang, Q.; Gao, J.; Cai, Y.; Yin, H.; Gao, Y.; Zhao, J.; Liu, L.; Huang, J. Development and application of benthic macroinvertebrate-based multimetric indices for the assessment of streams and rivers in the Taihu Basin, China. *Ecol. Indic.* **2015**, *48*, 649–659. [CrossRef]

59.  Allan, J.D. Landscapes and riverscapes: The influence of land use on stream ecosystems. *Annu. Rev. Ecol. Evol. Syst.* **2004**, *35*, 257–284. [CrossRef]