4

Assessing environmental conditions in rivers and streams with diatoms

R. JAN STEVENSON, YANGDONG PAN, AND HERMAN VAN DAM

4.1 Introduction

Assessments of environmental conditions in rivers and streams using diatoms have a long history in which two basic conceptual approaches emerged. First, based on the work of Kolkwitz and Marsson (1908), autecological indices were developed to infer levels of pollution based on the species composition of assemblages and the ecological preferences and tolerances of taxa (e.g. Butcher, 1947; Fjerdingstad, 1950; Zelinka and Marvan, 1961; Lowe 1974; Lange-Bertalot, 1979). Second, Patrick's early monitoring studies (Patrick, 1949; Patrick et al., 1954; Patrick and Strawbridge, 1963) relied primarily on diatom diversity as a general indicator of river health (i.e. ecological integrity), because species composition of assemblages varied seasonally and species diversity varied less. The conceptual differences in these two approaches really address two different goals for environmental assessments, one inferring pollution levels and the other determining biodiversity, a more valued ecological attribute (Stevenson, 2006). Thus, the concepts and tools for assessing ecosystem health and diagnosing causes of impairment in aquatic habitats, particularly rivers and streams, were established and developed between ~50 and 100 years ago.

Today, diatoms are being used to assess ecological conditions in streams and rivers around the world (Asai, 1996; Kelly et al., 1998; Wu, 1999; Lobo et al., 2004; Wang et al., 2005; Chessman et al., 2007; Taylor et al., 2007; Porter et al., 2008). They have become valuable elements in large-scale national and international assessment programs of the United States and Europe (e.g. Kelly et al. 2009a). Keeping up with the large number of papers being published in locally, nationally, and internationally recognized journals has become a challenge, but this review shows the emergence of a great diversity of

methods and findings, identifying regionally refined tools, and organizing the application of these tools in scientifically sound protocols for solving environmental problems. Although initially overwhelming, the diversity of methods and findings for diatom assessments produced over the last ~100 years can be organized by relating them to frameworks for environmental assessment and management. Using the correct diatom assessment tools for the correct reason during an assessment is critical for scientifically rigorous and effective support of environmental management.

The many advances in using diatoms and other algae for monitoring stream and river quality have been reviewed by Patrick (1973) and, more recently, by Stevenson and Lowe (1986), Rott (1991), Round (1991), Whitton et al. (1991), Coste et al. (1991), Whitton and Kelly (1995), Rosen (1995), Lowe and Pan (1996), Ector et al. (2004), and Stevenson and Smol (2002). In this chapter, we update our last review completed for the first edition of this book (Stevenson and Pan, 1999). First, we provide the foundation for how diatoms can be used in ecological assessments of rivers and streams and why they should be used. Second, we review the many characteristics of diatom assemblages that could be used in assessments and the methods used for sampling, sample analysis, and calculating indicators. Finally, we provide examples and discuss how a set of diatom indicators can be used in ecological assessments with different purposes for each indicator, such as establishing reference conditions and numeric water-quality criteria, assessing physical and chemical as well as biological condition, and diagnosing stressors of ecosystem services.

4.2 Rationale for using diatoms

Two fundamental questions need to be answered in almost all ecological assessments: "Is there a problem?" and "What is causing the problem?" (Stevenson et al., 2004a). Understanding the meaning of these questions, and how they will be asked
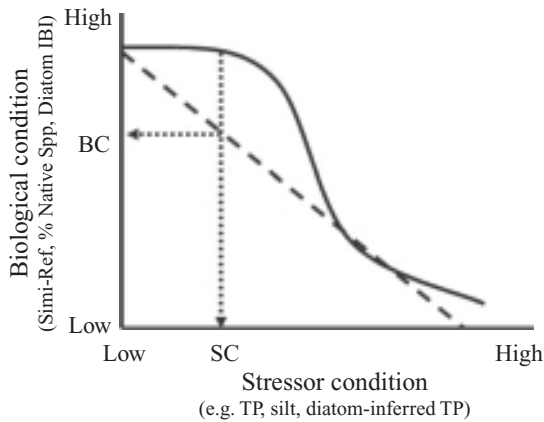
Figure 4.1 A scheme for using a suite of indicators for stressor and biocriteria development. The solid line represents a valued ecological attribute, such as the percentage of individuals in native species in the habitat, having a non-linear response to a stressor or human-disturbance gradient. The dashed line represents an indicator with linear response to a stressor or human-disturbance gradient. After identifying indicators with linear and non-linear responses, stressor criteria can be proposed for the highest stressor level (vertical dashed line) that has a high probability of protecting the valued ecological attribute with the non-linear response. The biological criterion (horizontal dashed line) can be proposed for an indicator with the linear response at a level predicted to occur when the stressor is at its criterion level.

and answered by government agencies or other scientists, is important for determining how diatoms can help answer both of these questions. Problems can usually be defined relative to the uses of water bodies and their ecosystem services. "Use" is a term that is defined in regulations developed in the United States of America (USA) associated with implementation of the US Clean Water Act (US CWA) and is one part of water quality standards. Examples of uses are aquatic-life use, drinking-water supply, and recreation. The other parts of water-quality standards in the USA are water-quality criteria, which are established to protect uses of waters, and an antidegradation policy that protects future degradation of waters. Water-quality criteria can be narrative or numeric descriptions of the physical, chemical, and biological condition of waters that define whether a use is being supported or not. The variables used for water-quality criteria are indicators of either uses or the contaminants and habitat alterations (i.e. "stressors") that degrade or threaten uses (Figure 4.1). Throughout this chapter, we will emphasize that slight differences in how diatom indicators are developed and applied determines whether indicators should be used to assess valued ecological attributes (uses and ecosystem services) or stressors (Stevenson, 2006).

Protection of the biota that we would expect in minimally disturbed waters is a common policy of governments around the world. In the USA, many state governments refer to this as "aquatic life use support" and measure it with indicators of biological condition (Davies and Jackson, 2006). Waters with high levels of biological condition have biological integrity, which is a defined goal of the US CWA. Biological condition is usually measured as some deviation from natural or minimally disturbed condition (Davies and Jackson, 2006; Stoddard *et al.* 2006). In the European Union (EU), all water bodies are supposed to achieve "good ecological status" by 2015 according to the Water Framework Directive (WFD: European Union, 2000). "Good ecological status" is defined in the EU WFD as biota consistent with only slight alterations from that expected in the absence of human impacts (reference conditions), which is the same as the concept of high biological condition in the USA. Although we usually measure structure, such as species composition of assemblages, both the US and EU concepts of reference conditions also include the function of aquatic ecosystems, such as productivity and nutrient uptake (Karr and Dudley, 1981).

One of the main reasons for convergence of management goals on "good ecological status" or "biological integrity" is that biodiversity is a valued ecological attribute, thus an ecosystem service. In addition, waters with high biological condition are assumed to support other ecosystem services, such as safe and aesthetically pleasing drinking water and recreational uses. In addition, rivers and streams are complex ecosystems in which many environmental factors vary on different spatial and temporal scales. Even though as many variables as possible should be measured to determine environmental conditions in a habitat (Barbour *et al.*, 1999; Norris and Norris, 1995), full assessment of all the potentially important variables is impractical. Biological indicators respond to altered physical and chemical conditions that may not have been measured and thereby reflect these altered conditions in assessments. Using a suite of biological indicators, based on organisms living from one day to several years, provides an integrated assessment of environmental conditions in streams and rivers which are spatially and temporally highly variable.

Using diatoms as indicators of environmental conditions in rivers and streams is important for three basic reasons: their importance in ecosystems, their utility as indicators of environmental conditions, and their ease of use. Diatom importance in river and stream ecosystems is based on their fundamental role in food webs (e.g., Mayer and Likens, 1987; see review by Lamberti, 1996) and linkage in biogeochemical cycles (e.g.

Newbold *et al.*, 1981; Kim *et al.*, 1990; Mulholland, 1996). As one of the most species-rich components of river and stream communities, diatoms are important elements of biodiversity and genetic resources in rivers and streams (Patrick, 1961). In addition, diatoms are the source of many nuisance algal problems, such as taste and odor impairment of drinking water, reducing water clarity, and clogging water filters (e.g. Palmer, 1962).

Diatoms are valuable indicators of environmental conditions in rivers and streams, because they respond directly and sensitively to many physical, chemical, and biological changes in river and stream ecosystems, such as temperature (Squires *et al.*, 1979; Descy and Mouvet, 1984), nutrient concentrations (Pan *et al.*, 1996; Kelly, 1998), and herbivory (Steinman *et al.*, 1987a; McCormick and Stevenson, 1989). The species-specific sensitivity of diatom physiology to many habitat conditions is manifested in the great variability in biomass and species composition of diatom assemblages in rivers and streams (e.g. Patrick, 1961). This great variability is the result of complex interactions among a variety of habitat conditions that differentially affect physiological performance of diatom species, and thereby, diatom assemblage composition. Stevenson (1997) organizes these factors into a hierarchical framework in which higher-level factors (e.g. climate and geology) can restrict effects of low-level factors. Low-level, proximate factors, such as resources (e.g. light, nitrogen (N), phosphorus (P)) and stressors (e.g. pH, temperature, toxic substances), directly affect diatoms. At higher spatial and temporal levels, effects of resources and stressors on diatom assemblages can be constrained by climate, geology, and land use (Biggs, 1995; Stevenson, 1997). The sensitivity of diatoms to so many habitat conditions can make them highly valuable indicators, particularly if effects of specific factors can be distinguished.

Diatoms occur in relatively diverse assemblages, and most species, especially common species, are relatively easily distinguished when compared to other algae and invertebrates that also have diverse assemblages. Diatoms are readily distinguished to species and subspecies levels based on unique morphological features, whereas many other algal classes have more than one stage in a life cycle and some of those stages are either highly variable ontogenically (e.g. blue-green algae), cannot be distinguished without special reproductive structures (e.g. Zygnematales), or cannot be distinguished without culturing (many unicellular green algae). Diverse assemblages provide more statistical power in inference models (see Birks, this volume). Identification to species level improves precision and accuracy of indicators that could arise from autecological variability within genera. Diatoms are relatively similar in size compared to the variability among all groups of algae, even though the size of diatoms does vary by several orders of magnitude; therefore diatom assemblage characterizations accounting for cell size (biovolume and relative biovolume) are probably not as necessary as when using all groups of algae together.

Diatoms can be found in almost all aquatic habitats, so that the same group of organisms can be used for comparison of streams, lakes, wetlands, oceans, estuaries, and even some ephemeral aquatic habitats. Diatoms can be found on substrata in streams, even when the stream is dry, so they can be sampled at most times of the year. If undisturbed sediments can be found in lotic ecosystems, such as in reservoirs, deltas, wetlands, and floodplains where rivers and streams drain, diatom frustules are preserved in sediments and record habitat history (Amoros and van Urk, 1989; Cremer *et al.*, 2004; Gell *et al.* 2005). Historical conditions in streams and rivers have also been assessed by using museum collections of diatoms on macrophytes and in fish (van Dam and Mertens, 1993; Rosati *et al.*, 2003; Yallop *et al.*, 2006). Diatoms have shorter generation times than fish and macroinvertebrates. Therefore we assume they reproduce and respond rapidly to environmental change, thereby providing early warning indicators of both pollution increases and habitat restoration success. The combined costs of sampling and sample assay are relatively low when compared to other organisms. Samples can be archived easily for long periods of time for future analysis and long-term records.

Thus, the study of diatoms has become an important element of monitoring and assessment programs in countries around the world. In the EU, each member state has established methods for assessing ecological status for a range of water-quality elements, including phytobenthos in running waters. Diatoms are used in water-quality assessments in many states in the USA and were included in the National Rivers and Streams Assessment being conducted from 2008–2011 by the US Environmental Protection Agency (EPA). Diatoms provide as precise an assessment, as macrophytes, invertebrates, and fish and actually are more sensitive to some stressors than other organisms (Griffith *et al.*, 2005; Johnson *et al.*, 2006; Chessman *et al.*, 2007; Carlisle *et al.*, 2008). Hering *et al.* (2006) state that diatoms are most sensitive to changes in nutrient and organic-matter concentrations, macrophytes are sensitive to changes in hydromorphology, and macroinvertebreates and fish are sensitive to changes in hydromorphology at larger scales of the landscape.

4.3 Assemblages and characteristics used in assessment

4.3.1 Periphyton, plankton, and sampling

Both planktonic and benthic diatoms (i.e. phytobenthos or periphyton) can be used in assessments of rivers and streams. The advantages of sampling plankton and periphyton vary with size of the river and objective of the research. Plankton should usually be sampled in large rivers and periphyton should be sampled in shallow streams, where each, respectively, is the most important source of primary production (Vannote *et al.*, 1980). However, periphyton sampling could be more appropriate than phytoplankton in large rivers if assessing point sources of pollution and if high spatial resolution in water-quality assessment are objectives. If greater spatial integration is desired, then plankton sampling may be an appropriate approach even in small streams. Suspended algae originate from benthic algae in small rivers and streams and are transported downstream (Swanson and Bachman, 1976; Müller-Haeckel and Håkansson, 1978; Stevenson and Peterson, 1991). Therefore, plankton may provide a good spatially integrated sample of benthic algae in a stream. Further studies of the value of sampling plankton in small rivers and streams are warranted for assessment of watershed conditions.

Benthic algae on natural substrata and plankton should be sampled in stream assessments whenever objectives call for accurate assessment of ecosystem components (Aloi, 1990; Cattaneo and Amireault, 1992) or when travel costs to sites are high. Artificial substratum sampling is expensive because it requires two separate trips to the field, and because artificial substrata are highly susceptible to vandalism and damage from floods. One problem with periphyton samples on natural substrata is that they can be highly variable. Targeted habitat sampling, such as rocks and hard surfaces, are recommended in the EU and some US programs (Kelly *et al.*, 1998; Moulton *et al.*, 2002; CEN, 2003, 2004). However, other US programs (Weilhoefer and Pan, 2007; Peck *et al.*, 2006) randomly sample whatever substratum occurs throughout a defined section of the stream to provide an accurate characterization of assemblages in that section (i.e. reach). Composite sampling approaches have been used to reduce within-habitat variability when sampling stream and river periphyton. Composite samples are collected by sampling periphyton on rocks at random locations along three or more random transects in a habitat and combining the samples into one composite (e.g. Weilhoefer and Pan, 2007).

Using artificial substrata is a valuable approach when objectives call for precise assessments in streams with highly variable habitat conditions, or when natural substrata are unsuitable for sampling. The latter may be the case in deep, channelized, or silty habitats. Benthic algal communities on artificial substrata are commonly different than those on natural substrata (e.g. Tuchman and Stevenson, 1980). However, when the ecology of the natural habitat is simulated, benthic diatom assemblages developing on artificial substrata can be similar to assemblages on natural substrata (see review by Cattaneo and Amireault, 1992). Cattaneo and Amireault recommend cautious use of artificial substrata, because algal quantity often differs and non-diatom algae are underrepresented on artificial substrata. When detecting change in water quality is a higher priority than assessing effects of water quality on natural assemblages of periphyton in that habitat, then the advantages of the high precision and sensitivity of diatoms on artificial substrata for assessing the physical and chemical conditions in the water may outweigh the disadvantages of questionable simulation of natural communities.

4.3.2 An organizational framework for assemblage attributes

To help organize the tools for the assessment toolbox, we have classified assemblage attributes based on how they are developed, because that affects how they should be applied. Diatom assemblage characteristics are typically used in conjunction with the characteristics of entire periphyton or plankton assemblages, thereby accounting for changes in other algae and microorganisms that occur in benthic and planktonic samples. These characteristics occur in two categories, structural and functional (Table 4.1). Taxonomic composition and biomass (measured as cell density or biovolume) are the only diatom assemblage characteristics that can be distinguished from other algae and microbes in periphyton and plankton samples, which can be done with microscopic examination, identification, and counting cells in samples. It is worthwhile to note that chlorophyll *a* (chl *a*) concentrations, ash-free dry mass (AFDM), chemical composition, and functional characteristics of diatom assemblages cannot be distinguished from other algae, bacteria, and fungi in a periphyton or plankton sample. Little is known about the accuracy of diatom biomass assessments with chl *c*, so measurements of diatom cell numbers or biovolumes are probably the most reliable estimates of diatom biomass.

The most commonly used diatom indicators are those using taxonomic composition of assemblages. There is a great diversity of these indices based on how they are calculated and how

Table 4.1 Characteristics of algal assemblages that could be used to assess the ecological (biological and stressor) condition of a habitat. Attribute type indicates the kind of analysis, whether attributes are structural or functional characteristics, and whether they are usually (not always) indicators of biological or stressor condition. For indicators and citations, we've provided early and recent examples, but have not attempted to be encyclopedic. (Abbreviations: SPI, Species Pollution Index; GDI, Indice Générique Diatomique; DAI, Diatom Autecological Index; TDI, Trophic Diatom Index)

Attribute type	Indicator	Citations
Biomass (structural characteristic, biological condition)	AFDM, chl *a* and other pigments, cell densities, cell biovolumes, other elements that are most common in microbial biomass (N or P)	APHA, 1998; Dodds *et al.* 1997; Stevenson *et al.* 2006
Diversity (structural characteristics, biological condition)	Composite diversity, total species or generic richness, richness of sensitive native taxa, species evenness, percent dominant taxon	Stevenson *et al.*, 1984b, 2008a
Taxonomic composition (structural characteristic, biological condition)	Similarity to reference condition, relative abundances of individuals in genera and species, pigment ratios	Common throughout the literature, e.g. Schoemann, 1976; Lange-Bertalot, 1979; Stevenson and Bahls, 1999; Passy and Bode 2004; Wang *et al.*, 2005; Lavoie *et al.*, 2006; Stevenson *et al.*, 2008a
Guild indicators (structural characteristics, biological condition)	Percent of diatom individuals or taxa that are stalked, adnate, motile, native (reference), non-native (non-reference), pollution sensitive, pollution tolerant or are adapted to low P, high P, low pH, high pH, etc.	Palmer, 1969; Lange-Bertalot, 1979; Wang *et al.*, 2005; Stevenson *et al.* 2008a, 2008b
Autecological indices (structural characteristic, stressor condition)	Pollution Tolerance Index, SPI, GDI, DAIpo, DAI-pH, DAI-TP, TDI, human disturbance, conductivity, pH, trophic status, saprobity (organic pollution), dissolved oxygen, nitrogen, phosphorus, sediments	Lange-Bertalot, 1979; Rumeau and Coste, 1988; Watanabe *et al.*, 1986; Prygiel and Coste, 1993; van Dam *et al.*, 1994; Kelly *et al.*, 1995; Pan *et al.*, 1996; Walley *et al.*, 2001; Gómez and Licursi, 2001; Rott *et al.*, 2003; Potapova and Charles, 2003, 2007; Lobo *et al.*, 2004; Kovács *et al.*, 2006; Stevenson *et al.*, 2008a, 2008b
Morphology (structural characteristic, biological & stressor condition)	Larger cells with UV effects, deformed frustules with metals	Bothwell *et al.*, 1993; McFarland *et al.*, 1997; Falasco *et al.*, 2009
Chemical ratios (structural characteristic, biological & stressor condition)	chl *a*:AFDM, chl *a*:phaeophytin, N:P, N:C, P:C, heavy metals:AFDM	Weber, 1973; Peterson and Stevenson, 1992; Humphrey and Stevenson, 1992; Biggs, 1995
Growth and dispersal rates (functional characteristic, biological condition)	Reproduction rate, growth rate, accrual rate, immigration rate, emigration rate	Müller-Haeckel and Hakansson, 1978; Stevenson, 1983; Biggs, 1990
Metabolic rates (functional characteristic, biological condition)	Photosynthetic rate, respiration rate, phosphatase activity	Blanck, 1985; Mulholland and Rosemond, 1992; Hill *et al.*, 1997

they should be applied. Often, two pieces of information are needed to calculate diatom indicators that use taxonomic composition of assemblages, the relative abundance of taxa in the sample (i.e. taxonomic composition) and the environmental preferences and tolerances of taxa (autecological attributes of taxa or taxa traits). These calculations are usually done at the species level, but they can be done at higher levels of taxonomic organization as well.

Taxa autecologies or traits (*sensu* Stevenson *et al.* 2008b) were historically determined by non-quantitative rankings based on the types of habitats in which taxa had been collected in the past. The compilations of literature with indicator values for diatoms by Lowe (1974) and van Dam *et al.* (1994) are two examples. The weighted-averaging method, developed and refined from 1955 to 1989 (Pantle and Buck, 1955; Zelinka and Marvan, 1961; ter Braak and van Dam, 1989), is a more quantitative method for determining species environmental optima and the tolerance of taxa for environmental variability. Morphological traits, such as stalked or adnate growth forms and raphe structures (indicating ability to move on substrata

and through a periphyton matrix and sediments), are attributes that can be used to group taxa and start identifying guilds of taxa, i.e. groups with similar functions in ecosystems (see Wang *et al.* 2005). Physiological guilds of species can also be defined with low or high trait values, such as: low and high nutrient taxa, native and non-native taxa, taxa sensitive and tolerant to pollution, or low pH and high pH taxa (e.g. van Dam *et al.*, 1994; Stevenson *et al.*, 2008a). Guilds can also be identified using indicator species analysis (Dufrêne and Legendre, 1997), which allows selection of species that are more consistently found in one type of habitat versus another. Alternatively, we can assume that taxa, such as genera, families, orders, and classes have greater physiological similarity within taxa than across taxa. Thereby, percent *Nitzschia*, araphids, and Centrales become indicators based on the assumption that their physiologies or performance in ecosystems differ from other taxa. Testing guild indicators can be as simple as a regression analysis (e.g. Stevenson *et al.*, 2008b). Alternatively, Potapova and Charles (2007) use indicator species analysis to identify taxa characteristically found in high and low total phosphorus (TP) or total nitrogen (TN) conditions.

Usually, diatom indicators are calculated based on a weighted-averaging model using relative abundances of all taxa in a sample from a site and the environmental optima or autecological values for the taxa. However, they can be calculated many other ways. One major distinction is using subsets of all taxa in indicators (such as guilds) versus all taxa. The coarsest resolution for indicators using multiple species would be the simple average of species traits (environmental optima or autecological values) for all species in the sample; thus all species are weighted equally rather than by their relative abundances (Stevenson *et al.*, 2008b). Other calculation methods include the number of taxa in one specific guild, the proportion of all taxa in samples from one specific guild, or the proportion of all individuals in samples from one specific guild (Wang *et al.*, 2005).

Diatom indicators based on taxonomic characteristics of assemblages and taxa traits have been used to assess both the biological condition of rivers and streams and causes of their degradation, thus physical, chemical, and biological condition. Traditionally, diatom assemblages have been used to assess the chemical condition, using saprobic, pH, nutrient, and other stressor indicator systems (e.g. van Dam *et al.*, 1994; Kelly and Whitton, 1995; Potapova and Charles, 2003; Walley *et al.*, 2001). More recently, diatom research has also emphasized assessment of the biological condition of an ecosystem, i.e. the similarity of an assessed assemblage to assemblages in

reference ecosystems (Passy and Bode, 2004; Wang *et al.*, 2005; Lavoie *et al.*, 2006; Stevenson, 2006).

There is a fine distinction between how these indicators are calculated, but there should be a substantial difference in how they are applied. This distinction is discussed extensively in Stevenson (2006). The basic point is we should distinguish conceptually between indicators of physical, chemical, and biological condition because they are used differently in management. They are used differently because of how we manage ecosystems; we want to manage ecosystems to minimize physical alteration and chemical pollution of habitats to maximize the biological condition (Karr and Dudley, 1981; Karr and Chu 1997; Davies and Jackson, 2006) or some other valued ecological attributes (Stevenson *et al.*, 2004a). Therefore, indicators such as similarity of species composition to reference condition, percent native and non-native taxa (i.e. reference or non-reference taxa), percent diatoms, percent *Achnanthidium* A. G. Agardh, and percent *Cymbella* Kutzing and *Encyonema* Kutzing are indicators of biological condition because they do not specifically infer abiotic conditions and they do reflect elements of biological condition (i.e. "the natural balance of flora and fauna . . . ", Karr and Dudley, 1981). Inferred TP, pH, and specific conductivity are indicators that reflect shifts in diatom species composition from reference condition, but they more specifically address abiotic conditions. Inferred abiotic conditions are not easily interpretable biological attributes. Indicators such as percent low P species or percent motile species are more easily interpreted and could be considered valued attributes if they were thought of as elements of biological condition, even though they are stressor-specific guilds.

4.3.3 Biomass assay

Periphyton and phytoplankton biomass can be estimated with assays of dry mass (DM), AFDM, chl *a*, cell densities, cell biovolumes, and cell surface area (Aloi, 1990; APHA, 1998). All these estimates have some bias in their measurement of algal biomass. Dry mass varies with the amount of inorganic as well as organic matter in samples. The AFDM varies with the amount of detritus as well as the amount of bacteria, fungi, microinvertebrates, and algae in the sample. Chl *a*:algal carbon (C) ratios can vary with light and N availability (Rosen and Lowe, 1984). Chl *c* concentration in habitats could be a good indicator of diatom biomass in a habitat (APHA, 1998). Cell density:algal C ratios vary with cell size and shape. Even cell volume:algal C ratios vary among species, particularly among some classes of algae, because vacuole size in algae varies (Sicko-Goad *et al.*, 1977). Cell surface area may be a

valuable estimate of algal biomass because most cytoplasm is adjacent to the cell wall. Elemental and chemical mass per unit area (other than chl *a*, such as µg P cm^{-2}, µg N cm^{-2}) could also be used to assess algal biomass, but many of the chemical methods of assessing biomass have not been studied extensively.

We recommend using as many indicators of algal biomass as possible. We typically do not restrict our assays of algal biomass to diatom density and biovolume. We usually assess chl *a* and AFDM of samples and count and identify all algae to the lowest possible taxonomic level in Palmer counting cells or wet mounts to determine algal cell density and biovolume. More recently we have employed rapid periphyton surveys to visually assess biomass of microalgae and macroalgae with substantial success (Stevenson and Bahls, 1999; Stevenson *et al.*, 2006). In ecological studies when distinguishing live and dead cells is important (e.g. experiments), diatoms are counted in syrup (Stevenson, 1984a) or high refractive-index media using vapor substitution (Stevenson and Stoermer, 1981; Crumpton, 1987). In large-scale ecological surveys, when distinguishing live and dead diatoms has not been shown to be important, we count acid-cleaned diatoms using a highly refractive mounting medium such as Naphrax® to ensure the best taxonomic assessments. Conceptually, counting dead diatoms that may have drifted into an area or persisted from the past should only increase the spatial and temporal scale of an ecological assessment. Gillett *et al.* (2009) show that indicators performed similarly if they were based on diatom frustules with protoplasm or all diatoms in acid-cleaned material.

Periphyton and phytoplankton biomass is highly variable in streams and rivers, and periphyton biomass in particular has been criticized as a reliable indicator of water quality (Whitton and Kelly, 1995; Leland, 1995). According to theories of community adaptation to stress (Stevenson 1997), biomass should be less sensitive than species composition to environmental stress, because communities can adapt to environmental stress by changing species composition.

Another problem with using algal biomass as an indicator of nutrient enrichment and toxicity is that low biomass may be the result of a recent natural physical and biotic disturbance (e.g. Tett *et al.*, 1978; Steinman *et al.*, 1987) or toxicity (e.g. Gale *et al.*, 1979). A more reliable indicator of environmental impacts on algal and diatom biomass in a habitat may be the peak biomass that can accumulate in a river or stream after a disturbance (Biggs, 1996; Stevenson, 1996). Peak biomass is the maximum biomass in the phytoplankton or the periphyton that accumulates after a disturbance. These maxima develop during low discharge periods, usually seasonally, for both phytoplankton and periphyton and, theoretically, should be highly correlated to nutrient and light availability in a system. Clear relationships have been shown between phosphate and nitrate concentration and peak biomass of periphytic diatoms in experimental systems (Bothwell, 1989; Rier and Stevenson, 2006). Peak biomass is also a valuable parameter because it indicates the potential for nuisance-levels of algal biomass accumulation. In practical application, we seldom sample at the period of peak biomass; however adjusting the timing of sampling to get as close as possible to peak biomass state should reduce variability in biomass relationships with nutrient concentrations. For river phytoplankton, consider use of remote sensing from satellites to assess algal biomass at large spatial scales.

More recent field studies have shown clear relationships between benthic algal biomass and nutrient concentrations in streams by either accounting for temporal variability with multiple measurements during a season (Stevenson *et al.*, 2006), time since last disturbance (Biggs, 2000), or by high sample size (Dodds *et al.*, 1997). In addition, Stevenson *et al.* (2006) show that response of biomass to nutrients is much more related to increases in filamentous green algae than diatoms; the response of diatom biomass is saturated at relatively low phosphorus concentrations. These relationships were characterized by fairly high variability, but also by a high magnitude of effect.

4.3.4 Diversity

Many indices have been developed to characterize the number of species in a sample (species richness), the evenness of species abundances, and composite diversity. Composite diversity is represented in indices that respond to changes in both richness and evenness (e.g. Shannon, 1948; Simpson, 1949). High correlation between all of these indices has been observed (Archibald, 1972). This is probably because composite diversity and species richness measurements are highly dependent on evenness of species abundances in short counts (e.g. 600 valves; Stevenson and Lowe, 1986).

Assessment of species richness in a habitat is particularly problematic because species numbers are highly correlated to species evenness in counts when a predetermined and low number of diatoms is counted (e.g. 600 valves). Better assessments of species richness can be determined by developing the relationship between species numbers observed and the number of cells that have been counted. Species richness can be defined as the number of species in a count when no new species are observed with a specified additional counting effort.

Alternatively, non-linear regression can be used to determine the asymptote of the relationship between number of species observed and number of cells counted – the asymptote is an estimate of the number of species in the sample. The precision of the asymptote, and thus the estimate of species richness, will be reported by most statistical programs. Stratified counting efforts can be employed to assess different community parameters, such as relative abundance of dominant taxa and species richness. This is done by identifying and counting all valves until a pre-specified number of valves is enumerated to determine relative abundance of the dominants; then species richness is estimated by continuing to scan the sample, counting valves, identifying only the new species, and stopping the count when no new species are identified after a fixed counting effort, such as 100 or 200 valves. When the budget permits, species richness could be determined with long diatom counts (3000–8000 valves) and estimation using the assumption that the number of species in different density categories fit a log-normal curve (Patrick *et al.*, 1954). Assessments of species evenness do vary with the evenness parameter used (e.g. Hurlbert's versus Alatalo's evenness, Hurlbert, 1971; Alatalo, 1981), but the utility of differential sensitivity of these characteristics has not be extensively investigated.

The best use of diversity-related indices in river and stream assessments is probably as an indicator of changes in species composition when comparing impacted and reference assemblages (Stevenson, 1984b; Jüttner *et al.*, 1996). Some investigators have found that diversity decreases with pollution (e.g. Rott and Pfister, 1988; Sonneman *et al.* 2001), that diversity can increase with pollution (e.g. Archibald, 1972; van Dam, 1982; Stevenson *et al.* 2008a), and that diversity changes differently depending upon the type of pollution (Jüttner *et al.*, 1996). Patrick (1973) hypothesized ambiguity in diversity assessments of pollution when using composite diversity indices because of differing effects of pollutants on species richness and evenness. She predicted that some pollutants (e.g. organic pollution) would differentially stimulate growth of some species and thereby decrease evenness of species abundances. The author also predicted that toxic pollution could increase evenness and that severe pollution could decrease species numbers (Patrick, 1973). Therefore, depending upon the kind and severity of pollution, human alteration of river and stream conditions could decrease or increase the diversity that was characterized with composite indices that incorporate both the richness and evenness elements of diversity.

More recently Stevenson *et al.* (2008a) constrained diversity to just the sensitive taxa found in reference sites. Even though the total number of species in 600 valve counts increased over the TP gradient in this set of streams, the number of low-P native taxa decreased. After taking into account concerns about underestimates of taxa numbers in samples when using 600 valve counts and effects of evenness on taxa numbers observed in 600 valve counts, it is difficult to draw firm conclusions about effects of enrichment on diatom diversity in streams. However, we know that low-nutrient streams are dominated by a few small species of *Achnanthidium* (Kawecka, 1993; Stevenson *et al.*, 2008a) and that increases in nutrient concentrations may stimulate growth rates of other taxa more than the small *Achnanthidium* (Manoylov and Stevenson, 2006). Therefore, evenness of species densities likely increases with modest levels of nutrient enrichment causing an apparent increase in species numbers in samples when based on 600 valve counts. Key unanswered questions include, "How do we define and detect species extinctions?" and "Under what conditions do we actually have local, regional, and global extinctions?"

4.3.5 Taxonomic similarity

Changes in species composition tend to be the most sensitive responses of diatoms and other microbes to environmental change (van Dam, 1982; Niederlehner and Cairns, 1994). However, the temporal scale of the observation is important. In the very short term of a bioassay, algal metabolism responds sensitively to environmental stress (Blanck 1985). Quickly, however, communities can adapt to many environmental stresses by changing species composition and, thereby, may achieve biomass and metabolic rates like those in unimpacted areas (Stevenson, 1997). Diatom assemblages in most field situations have had this time to adapt to moderate environmental stresses by changing species composition. Therefore, in most field sampling situations, when stresses have existed long enough for immigration of new species and accrual of rare taxa that are stress-tolerant, species composition should be more sensitive to changes in environmental conditions than changes in biomass or metabolic rates (e.g. Schindler, 1990).

Ordination, clustering, and community similarity indices are three approaches to assess variation in species composition among communities. Ordination (correspondence analysis, detrended correspondence analysis, non-metric multidimensional scaling) is typically used to assess the multidimensional pattern in the relationships between assemblages based on species composition. Species and sample scores are related to ordination axes and can be used to determine which species were most important in groups of samples. Environmental conditions can also be related to the ordination axes by using

canonical correspondence analysis and detrended canonical correspondence analysis (ter Braak and Šmilauer, 2002). Ordination and clustering can be used to show which assemblages are the most different from other assemblages, which may be caused by anthropogenic impacts (e.g. Chessman, 1986; Stevenson and White, 1995).

Community dissimilarity or similarity indices (see reviews in Wolda, 1981 and Pielou, 1984) can be used to test specific hypotheses about correlations between changes in species composition and the environment (Cairns and Kaesler, 1969; Moore and McIntire, 1977; Peterson and Stevenson, 1989, 1992; Kelly, 2001; Passy and Bode, 2004; Stevenson *et al.* 2008a). Similarity of species composition to reference condition is one of the key metrics of biological condition recommended in the USEPA's Rapid Bioassessment Protocols (Stevenson and Bahls, 1999). For example, Stevenson *et al.* (2008a) observed thresholds in community similarity to reference condition along TP gradients. This metric provides a simple, easy-to-understand description of the percent change in species relative abundances between assemblages at assessed and reference sites. Lavoie *et al.* (2006) used the dominant axes in correspondence analysis to measure dissimilarity among communities along a human disturbance gradient, but the interpretation of this metric can be challenging for non-scientists. Cluster analysis (e.g. TWINSPAN, Hill, 1979) groups assemblages based on the similarity in species composition between assemblages (Leland, 1995). Community dissimilarity indices can also be used to distinguish among groups of assemblages by testing the hypothesis that dissimilarity among assemblages within a group is significantly less than between groups. Discriminant analysis can also be used to determine whether species composition of groups of assemblages differ significantly between clusters (ter Braak, 1986; Peterson and Stevenson, 1989). Multivariate analyses of dissimilarity among groups of assemblages is available in many modern statistical packages (e.g. McCune and Grace, 2002; Oksanen, 2004; Clarke and Gorley, 2006).

Taxonomic similarity among assemblages can also be evaluated with subsets of the data, such as percent Centrales, percent araphids, and percent *Nitzschia* Hass. Wang *et al.* (2005) tested the proportion of species and individuals in assemblages that were within one genus as potential metrics of biological condition. Many genera responded predictably to one environmental gradient or another. Percent *Cymbella* (*sensu lato*) individuals was an example of a taxon that has been adopted for use as a metric in some assessments (KDOW 2008). The transferability of genus level metrics from one study to another may

be problematic because of the variability in species that may occur across seasons or regions. Although these metrics do not explicitly involve morphological or physiological traits, they do implicitly; they will only be transferable if the function of the diatoms in ecosystems is genetically regulated by traits that vary less within genera than between genera. Thus they are similar to the following group of indicators, the guild indicators.

4.3.6 Guild indicators

Changes in relative abundance of indicator taxa or guilds are another type of measure of biological or stressor condition based on taxonomic composition, but they are also based on taxa traits. We will refer to this type of indicator as a guild indicator because these measures are based on subsets of the all taxa having similar physiologies and functions in the ecosystem. They differ from the weighted-average style indices because the latter use all taxa in the indicator calculation. In streams and rivers, the ratio between centrics and pennates is not a commonly used indicator, but centric and pennate diatoms are examples of indicator taxa at the class level of taxonomy (Brugam and Patterson, 1983). This ratio could be a valuable indicator because these classes of diatoms have different functions in streams and rivers, as centrics are primarily adapted to planktonic habitats and most pennates are benthic. Therefore a centric to pennate ratio could indicate both biological condition and hydrologic alteration by impoundments (a stressor in the environment) (Brugam *et al.* 1998).

Many guilds are genera or groups of genera, and most genus level indicators are at least implicitly based on the indicator taxon or guild concept. For example, the number of *Nitzschia* (*sensu lato*) taxa and relative abundance of *Nitzschia* taxa or individuals in samples are examples of guild indicators in which *Nitzschia* is the indicator taxon as well as the guild. This is because *Nitzschia* is often assumed to be a group of taxa that tend to be tolerant to a variety of pollutants (e.g. Wang *et al.*, 2005). The percentage of *Eunotia* Ehr. in a habitat can be used to infer the pH of habitats, particularly the relative pH of two habitats when diatom assemblages are compared. The dominance of a few small *Achnanthidium* taxa in low nutrient streams, because they grow faster than other taxa in low nutrients (Manoylov and Stevenson 2006), underpins the value of two indices: the ratio between percent *Achnanthes* Bory, *Cocconeis* Ehr. plus *Cymbella* versus percent *Cyclotella* (Kütz) Bréb., *Melosira* C. A. Agardh plus *Nitzschia* (Wu, 1999; note here *Achnanthes*, *Cocconeis*, *Cyclotella*, and *Melosira* were *sensu lato*, and *Achnanthes sensu lato* and *Achnanthidium* are the same in most freshwater cases); and the ratio

between percent *Achnanthidium* versus percent *Achnanthidium* plus *Navicula* Bory (Wang *et al.*, 2005).

Diatoms with the morphological capability of moving through fine sediments (e.g. *Cylindrotheca* Rab., *Gyrosigma* Hass., *Navicula*, *Nitzschia*, and *Surirella* Turp. in Wang *et al.*, 2005 and modified from Bahls, 1993) represent another guild in the community. Motile diatoms reflect an important function in habitats where fine sediments occur naturally, and therefore also reflect an element of biological condition. Motile diatoms can also be used as an indicator of excess siltation, one of the most widespread stressor conditions in streams and rivers. *Epithemia* Kütz. and *Rhopalodia* Müll. represent another guild of diatoms because they can have N-fixing blue-green endosymbionts (DeYoe *et al.*, 1992), and they respond when N occurs in relatively low supply compared to P in streams (Peterson and Grimm, 1992; Kelly, 2003); thus the percentage of these two genera in assemblages could be used as an indicator of low N concentrations or low N:P ratios. Growth forms could also be defined as a guild and tested as indicators of ecological condition (Wang *et al.*, 2005). In general, genus level metrics are not as precise as species level indicators, but they are valuable when only genus level taxonomy is available (Coste *et al.*, 1991; Prygiel and Coste, 1993; Kelly *et al.*, 1995; Hill *et al.*, 2001). In addition, guild indicators that usually use less than half the taxa in an assemblage are usually not quite as precise and transferable as weighted-average autecological indices that use all taxa. However, guild indicators can be more clearly interpreted for some applications than indicators using all taxa.

4.3.7 Weighted average autecological indices

Weighted average autecological indices are defined here as indices using traits of species. Historically, traits of species were defined categorically on an ordinal scale (e.g. Lowe 1974, van Dam *et al.* 1994, Stevenson *et al.* 2008a). Much trait information in the literature (e.g., Hustedt, 1957; Cholnoky, 1968; Slàdecek, 1973; Lowe, 1974; Descy, 1979; Lange-Bertalot, 1979; Beaver, 1981; Fabri and Leclercq, 1984; Steinberg and Schiefele, 1988; van Dam *et al.*, 1994; Gómez and Licursi, 2001; Rott *et al.*, 2003; Lobo *et al.*, 2004; and see reviews in Lowe 1974; van Dam *et al.*, 1994; Kelly and Whitton, 1995; Stevenson *et al.* 2008b) uses an ordinal scale for the species traits with six or less ranks for a specific environmental stressor (pH, nutrient requirements, temperature, salinity, organic pollution, etc.). See van Dam *et al.* (1994) for additional references. Remarkably precise autecological indices of environmental conditions can be calculated with the ordinally ranked traits using this great wealth of autecological information in the literature and

a simple formula (Pantel and Buck, 1955; Zelinka and Marvan, 1961):

$$\sum p_i \Theta_i / \sum p_i$$

in which p_i is the proportional abundance of the ith taxon for which autecological information is known (p_i, for $i = 1, 2, \ldots S$ species) and Θ_i is the autecological rank of a species for a specific stressor. These autecological indices tend to work across broad geographic ranges when more accurate information is not known about diatom traits in the region studied (Kelly *et al.*, 2009a; Porter *et al.*, 2008; Lavoie *et al.*, 2009; Stevenson *et al.*, 2009).

More recently, weighted-average optima for taxa have been developed extensively for many taxa, for many environmental variables, and in many regions (Table 4.1). Optima and sometimes tolerances have been developed in many regions for nutrient concentrations (Bennion *et al.*, 1996; Pan *et al.*, 1996; Schönfelder *et al*, 2002; Soininen and Niemelä, 2002; Potapova and Charles, 2007; Ponander *et al.* 2007). The tolerance of species to environmental variability can be used to downweight species influence on stressor inference models in a slightly modified version of the Zelinka and Marvan equation:

$$\sum p_i v_i t_i / \sum p_i t_i$$

in which v_i is the optimum environmental condition for a taxon and t_i is the tolerance of species to variation in environmental conditions (ter Braak and van Dam, 1989). Optima have also been developed for pH and conductivity gradients (Pan *et al.*, 1996; Potapova and Charles 2003). Stevenson *et al.* (2008b) developed optima for species performance along a human disturbance gradient for streams in the western USA. In this paper, they also document the importance of linking causality to indicators in regions when many environmental factors covary. Caution should be exercised when developing weighted-average models for stressors in large-scale data sets when environmental factors covary. Resulting inference models may not be robust over time and may not have a causal foundation, i.e. truly respond to the correlated physicochemical variable. Weighted-average optima have the advantage over ordinally ranked traits because they are continuous variables and may provide more precision than autecological indices that use categorical traits or guild indicators using subsets of species (Stevenson *et al.* 2008b).

Diatom-based autecological indices can be particularly valuable in stream and river assessments because one-time assay of species composition of diatom assemblages in streams could provide better characterizations of physical and chemical
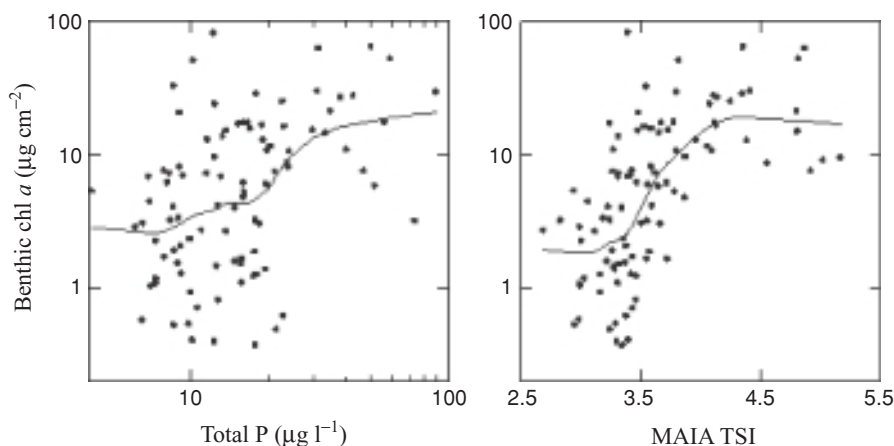
Figure 4.2 Refining stressor–response relationships with diatom indicators (Stevenson, unpublished data). The relationship between benthic algal biomass (chlorophyll *a*, $\mu g\ cm^{-2}$) and either total phosphorus concentrations or a trophic status index (Stevenson *et al.* 2006) in streams of Michigan and Kentucky, USA. Data for the same streams are plotted in both figures. The lines through the data show statistically significant linear and non-linear trends. The Mid-Atlantic Integrated Assessment Trophic Status Index (MAIA TSI) was developed in Stevenson *et al.* (2008a).

conditions than one-time measurement of those conditions. Precise characterization of environmental conditions in rivers and streams is difficult because of the high variability in discharge, water chemistry, temperature, and light availability associated with weather-related events. Charles (1985) showed that diatom-inferred pH was a better characterization of mean annual pH than one-time sampling of pH for lakes. The RMSE for a weighted-average index of TP based on diatom species composition in streams of the Mid-Atlantic Highlands (USA) was 0.32 log(TP $\mu g\ l^{-1}$) (Pan *et al.*, 1996) and was substantially less than the range in TP concentration that is commonly observed in streams. For example, TP ranged from 1.0–4.0 log(TP $\mu g\ l^{-1}$) over an eight week period in Kentucky streams (Stevenson, unpublished data). Although additional results from this study are not yet published, they provide evidence that diatom-inferred TP concentrations can more accurately assess bioavailable phosphorus than even repeated measurements of TP. Stevenson (unpublished data) shows that the conditions in which *Cladophora* accumulates across ecoregions can be better predicted with diatom-inferred TP than average TP measured over a two-month period (Figure 4.2). These results show how diatom-inference models for TP can aid refinement of stressor–response relationships, development of nutrient criteria to prevent nuisance algal blooms, and complement measured TP in use as nutrient criteria (see later discussion).

Variations on the weighted-average modeling approach have also been explored and provide some improvements in performance under specific circumstances. One of the most useful is weighted-average partial least squares (WA-PLS) technique, which builds a complex regression model that predicts environmental conditions based on species composition and regression coefficients for each species (ter Braak and Juggins, 1993). Indeed, Stevenson and Pinowska (2007) found WA-PLS to be a valuable indicator development method for nutrient models in Florida Springs, where variability in species composition was likely less than in broad-scale regional studies. In the northern Piedmont area of the eastern USA, Potapova *et al.* (2004) also document the utility of WA-PLS models for nutrients. In addition, Potapova *et al.* revisit the concept used by Zelinka and Marvan (1961) in which multiple indicator values are used per species and indicator system, such that each species is assigned a probability of being observed in a specific level of ecological condition, e.g. five successively higher TP ranges along a TP gradient (Potapova *et al.*, 2004). The authors conclude that despite the non-symmetrical distributions of many diatom taxa along the TP gradient, indicators perform just as well if they are based on the simplifying assumption that species have a symmetrical distribution along environmental gradients.

4.3.8 Sources of error

Considerable debate exists over sources of error in diatom assessments and how to reduce them. What habitats should be sampled and how many samples should be collected at a site? How many diatoms should be counted? How important is the level of taxonomic resolution and variability in taxonomy among operators/counters/technicians? The answer to these questions again depends on the objectives and budget of the project and indicator metrics.

First, the sampling substrata should be standardized if possible. Although some studies show that sampling multiple habitats and differing habitats among streams has relatively little effect on indicators and assessments of ecological conditions (Winter and Duthie, 2000; Potapova and Charles, 2005; Weilhoefer and Pan, 2007), others show or recommend that substrata sampled should be standardized as much as possible (Kelly *et al.*, 1998; Besse-Lototskaya *et al.*, 2006). Given that all streams do not have the preferred hard substrata for sampling, one reasonable approach is to separate stream types in analyses to have different *a-priori* defined reference conditions for low-gradient streams, in which there are few "hard" substrata, and higher-gradient streams with hard substrata.

Another sampling consideration is the number of samples that should be collected at a site. The answer to this issue depends upon the objectives of the study and definition of "samples." First, almost all protocols call for collecting multiple samples at a site during each visit and combining those samples into one or more composite samples that will be taken to the laboratory for separate analysis. For example, the USEPA protocols (Lazorchak *et al.*, 1998; Peck *et al.*, 2006; Stevenson *et al.*, 2008b) require sampling of substrata at randomly selected locations along each of nine transects at a site and combining those samples into a composite sample for analysis in the laboratory. It is assumed that this form of composite sampling reduces variation in assessments of a site. Taking one sample from a site is routine for determining the status of water bodies in a region when many water bodies are sampled or for development of relationships that can be used for development of water-quality criteria. If multiple composite samples were collected from a site, then only one of the composite samples could be used in statistics for regional assessments or developing relationships between ecological condition and stressors because replicate samples from an site are not statistically independent (Hurlbert, 1984). Alternatively advanced statistical methods can be used to account for lack of statistical independence for repeated measurements from a site.

However, if the objective of the study were to assess one specific site and compare it to standards, then multiple composite samples should be collected from one site to establish the central tendency and variation in assessed conditions. Kelly *et al.* (2009b) show that increased replication of samples, in this case taken at least six times over multiple years, reduces risk of misclassification of sites. They show that assessed conditions are most variable at intermediate levels of ecological condition. They recommend up to six samples at separate times for assessments of sites, with number of samples increasing if assessed conditions are close to those established as water-quality criteria. These considerations are particularly applicable to situations where individual sites are suspected of failing minimum standards, are candidates for restoration, or any situation when assessment of a specific site is needed.

Laboratory methods, as well as field-sampling methods, should be considered to minimize errors in assessments. The recommended number of valves to count in a sample ranges from 200 to tens of thousands, depending upon objectives. If an objective of the study is to determine the number and identity of most of the species in an assemblage, then 3000–8000 or even more cells may need to be counted (Patrick *et al.*, 1954). Alternatively, if a precise characterization of just the dominant taxa is necessary, then between 500 and 1000 diatoms should be counted, depending upon the number and evenness of species in the community (Pappas and Stoermer, 1996). The European Standard (CEN 2004) for the WFD has to be followed for identification and enumeration of diatoms. A typical count size is 300 to 500 valves, although lower or higher numbers may be appropriate for some purposes. Prygiel *et al.* (2002) show little effect for calculating the biological diatom index if more than 300 frustules or valves are counted. Routinely, 600 valves of diatoms are counted in the large-scale national programs in the USA (Charles *et al.*, 2002; Stevenson *et al.*, 2008b). One rationale for that count size is that 30–50 valves may need to be counted to estimate precisely the relative abundance of a taxon (Alverson *et al.*, 2003). Probably the best advice is to err on the high side of counting unless specific indices and tests of those indices have been calculated. Ultimately, we should be counting each sample until a desired level of precision in metrics is achieved. This varies with diversity and other factors, but will be easily incorporated into assessment protocols when software is developed so that analysts enter counts directly into the computer and have instantaneous updates of metric values as they count.

With species-level indicators, inter-analyst error in diatom identifications is an important source of variation in assessments (Prygiel *et al.*, 2002; Besse-Lototskya *et al.* 2006). Prygiel *et al.* (2002) found that misidentifications of small *Achnanthes* and *Cocconeis placentula* varieties were major sources of error for the Indice Biologique Diatomées (IBD) scale, and they recommend intercomparison exercises and internet exchange of materials to improve inter-analyst errors in taxonomy. Kahlert *et al.* (2009) show in a EU calibration test that the main sources of error were wrong calibration scales, overlooking small taxa (especially small *Navicula*), misidentifications (e.g. *Eunotia rhomboidea* was mistaken for *E. incisa*), and unclear separation

between certain taxa in the identification literature. In addition, Kahlert *et al.* (2009) show harmonization (frequent communication about taxonomy among analysts) is even more important than many years of experience in getting reliable monitoring results. Harmonization efforts with annual meetings, frequent internet communication, and a standardized taxa list have been important elements of taxonomic consistency in diatom programs in the USA, but results of these efforts have not been evaluated. In the USA, a standardized list of taxa names, with literature citations, and a type image was established for the US Geological Survey (USGS) National Water Quality Assessment (NAWQA) program. This protocol with a standardized list and taxa references is being followed in the National Lakes Survey and National Rivers and Streams Assessment sponsored by the USEPA.

One possible solution for the problem with taxonomic consistency is to use genus-level indicators. Although such indicators have significant value (Kelly *et al.*, 1995; Hill *et al.*, 2001; Wunsam *et al.*, 2002), they do not work as reliably in different regions and across larger geographic scales as species-level indicators (Chessman *et al.*, 1999; see early discussion of genus level indicators under headings for taxonomic and guild indicators).

Regionally varying factors and cryptic species can also be sources of error in all types of diatom indices. Potapova and Charles (2002, 2007) show that refining diatom optima for different regions of the USA improves performance of diatom indicators. Stevenson *et al.* (2008b), however, show that diatom indices are most precise when using data from multiple regions; but this may be due to increasing the length of environmental gradients with multiple regions in analyses. Regional variation in population traits, biogeography of taxa, interactions among populations present, or physical and chemical determinants of diatom ecology could affect transferability of metrics from one region to another. Biogeographic patterns in species, endemic species, and cryptic species indicate global distribution of populations is not sufficiently great to homogenize populations. Therefore, species traits may vary geographically. In addition, variation in species within genera could affect genus and guild indicator performance in the same way that variation in population genetics among regions affects species-level indicators. We will discuss ways to account for regional and between-stream variation in reference condition in a later section on defining reference condition. Thus, testing indicators in the regions that they will be used is important for evaluating their precision and accuracy, i.e. "Can they detect differences?" and "Are they detecting the differences we think they are?"

4.3.9 Robust indicators

Given all these sources of error, why are diatom indicators so robust? Why can we use them in multiple regions and incorporate results of many technicians in an assessment? It is generally agreed that regional refinements of indicators can improve indicator performance, and greater taxonomic consistency among analysts will improve assessments. But diatom indicators seem to be greatly transferable in space and time given the natural, sampling, and inter-analyst sources of variability. For example, the indicator values of van Dam *et al.* (1994) have been used in ecological assessments and indicator refinements in the USA (e.g. Porter *et al.*, 2008; Lavoie *et al.*, 2009; Stevenson *et al.*, 2009). Surely there was great variability in sampling and inter-analyst taxonomy between the sources of the indicator values in van Dam *et al.* (1994) and the cited projects. Pan (personal communication) has referred to this robustness of diatom indicators to sources of error as diatom-indicator magic.

Since this property of diatom indicators has not been evaluated in detail, we decided to postulate that the large number of taxa in an assemblage, unbiased error, and the Law of Large Numbers might explain why some types of diatom indicators are so robust. The Law of Large Numbers describes the stable behavior of the sampling mean of a random variable; such that repeated sampling from the population of a random variable with a finite expected mean will produce an estimated mean that approaches and stays close to the finite expected mean, when the sample size is large. Just as with flipping a coin repeatedly, the more times you flip the coin the closer you get to having 50% heads and 50% tails. The more diatoms you count, the closer you get to a stable mean of their relative abundances (Alverson *et al.*, 2003).

Let us start with a simple example to illustrate the concept, an assemblage in which the relative abundances of all species are equal. If we repeatedly sample species traits from the sample, then our estimate of the average trait value in the assemblage increases with the number of species observed from the assemblage. In addition, errors due to assigning traits to individuals or species are probably unbiased when they are related to biogeography, sampling, and inter-analyst taxonomic errors. Thus, with increasing error due to these sources, our assessments become less precise, but probably not less accurate. In addition, longer counts with more taxa and individuals should be more robust than shorter counts.

The manifestation of these properties for diatom metrics can be observed in the analyses conducted by Kahlert *et al.* (2009) during their analysis of taxonomic inconsistency among

analysts in northern Europe. Even though there was considerable variation in taxonomic identifications of some taxa among analysts, these errors seemed to be muted after calculation of the metrics. In Figures 1 and 2 of Kahlert *et al.* (2009), there is clear distinction between values of the Indice de Polluo-sensibilité Spécifique (Coste, 1982) and Acidity Index for Diatoms (Andrén and Jarlman, 2008) metrics among samples, despite the variation in taxonomic identifications of the analysts.

Our experience indicates that the robustness of diatom indicators seems to be greatest for weighted-average metrics. Indicators derived from cluster analyses and ordinations are more likely to be sensitive to taxonomic inconsistencies, because they do not involve an average of traits. The properties of diatom indicators should be studied more thoroughly so that we can understand how to improve their performance and which are more sensitive than others to different types of errors. Without question, we should strive to improve taxonomic consistency and accuracy of analysts, which will improve performance of all types of indicators. In addition, we should develop better traits for species as well as an understanding of the evolution of species, their adaptation to environments, and function in ecosystems. The collections of thousands of samples with complementary environmental information in assessment programs around the world provide an unprecedented opportunity for advancing these topics.

4.3.10 Morphological characteristics

Little research has been conducted to evaluate the effects of stressors on diatom size, striae density, shape, and other morphological characteristics. Sexual reproduction and auxospore formation in high-density periphyton assemblages after substantial colonization was hypothesized to be related to lower nutrient availability, in this case resulting from nutrient depletion developing during colonization and return of stream to baseflow conditions (Stevenson, 1990). Ultraviolet (UV) radiation may cause an increase in cell size and abundance of stalked diatoms (Bothwell *et al.*, 1993). Aberrant diatom shape, such as indentations and unusual bending in frustules, has been shown to be related to heavy-metal stress and many other stressors in streams (McFarland *et al.*, 1997). Falasco *et al.* (2009) provide an up-to-date review of the potential causes of different teratological forms for different species and specific stressors. Cells size, striae density, and shape of diatoms may also respond to environmental conditions. More research is justified to pursue this potentially sensitive set of metrics that could assist assessments of stressor as well as biological condition.

4.3.11 Chemical characteristics

Sediments and periphyton are important sinks for nutrients as well as many toxic inorganic and organic chemicals (Kelly and Whitton, 1989; Genter, 1996; Hoagland *et al.*, 1996). In addition to their potential as indicators of biomass, chemical characteristics of periphyton may provide valuable indications of the environmental conditions that affect periphytic diatoms. The TN and TP of periphyton communities have been used by Biggs (1995) to infer nutrient limitation and eutrophication in habitats (see also Humphrey and Stevenson, 1992). Kelly and Whitton (1989) demonstrate the accumulation of heavy metals in periphyton. Similarly, assays of particulate chemicals of the water column may provide insight into the chemical environment of phytoplankton that is not evident from assay of dissolved chemicals alone. For example, TP is used routinely to assess trophic status of lakes (Vollenweider and Kerekes, 1981). Thus chemical characteristics of assemblages could be used to assess biological and stressor condition of assemblages, but these indicators have not been fully developed and tested at broad scales.

4.3.12 Functional characteristics

Functional characteristics of diatom-dominated assemblages, such as photosynthesis and respiration rates, nutrient uptake rate and spiralling, phosphatase activity, and growth rate, have been used as indicators of environmental conditions in streams and rivers. Phosphatase activity is a valuable indicator of P limitation (Healey and Henzel, 1979; Mulholland and Rosemond, 1992). Photosynthesis and respiration can be used as measures of community productivity and health, but these assays are not commonly used in field surveys. Hill *et al.* (1997) use the response of periphyton respiration rate to experimentally manipulated stressors as an indicator of those stressors in the habitat. Since assemblages can adapt to environmental stressors by changing species composition and maintaining functional ecological integrity (Stevenson, 1997), Hill *et al.* (1997) predict that respiration rates of assemblages will not be inhibited by exposure to that stressor. Based on regional-scale patterns in phosphatase and respiration activity varying with TP concentration and water chemistry (Hill *et al.*, 1998; Stevenson *et al.* 2008a), metabolic indicators could be valuable in large-scale assessments of streams and rivers.

Growth rate has been used as an indicator of algal biomass production and can be assessed at population as well as assemblage levels (Stevenson, 1996). Schoeman (1976) and Biggs (1990) used growth rates as an indicator of nutrient limitation in habitats by resampling habitats after a short time

(3–7 days). Assessment of differing responses of species growth rates to environmental conditions may enhance the simple characterization of the autecology of species based on changes in their relative abundance. The main reason is that abundance is a function both of immigration and reproduction, as well as other processes (Stevenson, 1996). Reproduction is probably much more directly responsive to local conditions in a stream or river than immigration rates. However, assessing species growth rates in the field for assessments is impractical in most cases because of the need for multiple visits and perhaps use of artificial substrata. Assessing species growth rates in experiments and relating the data to field traits can be valuable. Manoylov and Stevenson (2006) link species responses to nutrients in experiments to the relative abundances in large-scale surveys, which helps to establish cause–effect relationships.

Thus, it is possible to measure ecosystem function in assessments of ecological condition, develop relationships between function and stressors, and thereby establish criteria for management. However, functional attributes are rarely measured in water-body assessments because the time needed for functional measures is greater than collecting materials in the field and later analysis of structural attributes in the laboratory. In addition, variability in ecological relationships is usually greater with functional attributes than many structural attributes, such as metrics based on species composition (Stevenson *et al.*, 2008a). Assessments of ecosystem function could be improved by better controlling for algal biomass. Metabolic rates are affected greatly by algal biomass (μg chl a cm^{-2}), whether metabolic rates are normalized based on area or biomass. Area-specific metabolic rates (e.g. mg C m^{-2} h^{-1}) increase with algal biomass, but biomass-specific metabolic rates (mg C μg^{-1} chl a h^{-1}) decrease with algal biomass (Stevenson, 1990; Hill and Boston, 1991). Future research should quantify the biomass effect on metabolism so that functional attributes of algal assemblages can be assessed.

Structural attributes can also be used to infer function. Numbers and biomass of functional groups of algae, invertebrates, and fish indicate the function of algae, grazing, and predation. Kelly *et al.* (2008) related shifts in diatom species composition to inferred changes in periphyton function. Trophic-state indices indicate nutrient conditions and could more specifically be related to algal biomass, productivity, and nutrient uptake. Future research should more directly address relationships between structure and function (similar to pattern and process) to confirm the often-implied assumption that water bodies with similar species composition and biomass have similar function, and to develop predictive models of ecosystem function.

4.3.13 Multimetric indices

Managing stream and river ecosystems calls for an assessment of integrity of the ecosystem and a diagnosis of causes of degradation (Figure 4.1). Indices of biotic integrity (IBI) of aquatic invertebrates and fish are being widely used to characterize streams (Karr, 1981; Hilsenhoff, 1988; Plafkin *et al.*, 1989; Lenat, 1993). These indices are often called multimetric indices of biological condition (IBC), because we are really assessing condition; integrity refers to high levels of biological condition. More recently, these multimetric indices (MMIs) have been developed for diatoms (Hill *et al.*, 2000; Fore and Grafe, 2002; Wang *et al.*, 2005; Kelly *et al.*, 2009a). The value of multimetric indices is they tend to be more precise than univariate metrics and they tend to be more linear (Fore *et al.*, 1994). In addition, they help provide a summary index and simplify communication of results. Thus, these indices are probably more valuable for assessing biological condition than non-linear univariate indicators because they respond sensitively to environmental change all along the human disturbance gradient (Figure 4.1).

Multimetric indices for diatoms can be constructed many ways, depending upon their goals. The basic steps for developing MMIs have been outlined by Plafkin *et al.* (1989) and more recently by Barbour *et al.* (1999), Wang *et al.* (2005), and Stoddard *et al.* (2008). First, the goal for the multimetric index should be established, such as an index for biological condition, stressor condition, or overall ecological condition (which would include biological and stressor conditions). A list of possible metrics should be developed based on the goal. Then the metrics should be classified for different elements of biological or stressor condition. For biological condition, examples of indicator classes are diversity, similarity to reference condition, sensitive and tolerant taxa, functional group, habitat, or growth form (Wang *et al.*, 2005; Stoddard *et al.*, 2008). For stressor condition, metrics should be considered for nutrient enrichment, siltation, acidification/alkalization, toxic substances, and hydrologic alteration. Both biological- and stressor-condition metrics would be included in an MMI for ecological condition. Metrics should then be adjusted for effects of natural variability and tested for having adequate range, reproducibility, and responsiveness (Stoddard *et al.*, 2008). Then the metrics with the highest responsiveness should be selected from each category in an iterative process that minimizes correlations with metrics from other categories. Metrics need to be rescored so they are all on the same scale, e.g. 0–1

or 0–100. Four to ten metrics are commonly used in MMIs for periphyton and diatoms (e.g. Hill *et al.*, 2000; Schaumburg *et al.*, 2004; Wang *et al.*, 2005; KDOW, 2008). Most assessment processes use the average of metrics in MMIs to evaluate the condition, but other decision processes could apply rules with joint criteria involving individual assessment of metrics with summarization in MMIs (see discussion in Wang *et al.*, 2005). An example of a rule with joint criteria is that all metrics must pass individual tests or a site fails to meet the criteria.

4.4 Assessment of biological condition and diagnosis of environmental stressors

4.4.1 Assessing ecological condition

Here we define ecological condition to include physical, chemical, and biological condition, and ecological integrity to be a high level of ecological condition. As emphasized previously, with diatoms we can measure biological condition of diatoms and infer chemical and physical condition, e.g. pH, nutrient concentrations, conductivity, and temperature. Ecological condition can be assessed as the absolute value of an attribute or as a deviation from an expected condition. For example, we could describe the number of pollution-sensitive species in a habitat, or we could describe the ratio between the observed and expected number of pollution-sensitive species in a habitat. Simple use of the number without comparison to an expected condition limits interpretation of the meaning. For example, when expected condition is a reference condition, then deviation from the reference condition can more clearly describe the effects of human activities. When the expected conditions are indicator values used as ecological status boundaries or water-quality criteria in the EU and USA, respectively, then they are triggers for management actions such as development and implementation of restoration plans.

4.4.1.1 Reference conditions and assessment
Minimally disturbed conditions have been defined using two basic methods, frequency distributions and predictive models. Assuming that we have an indicator that is highest in reference conditions (e.g. number of pollution-sensitive species), the 25th percentile of a frequency distribution of all sites or the 75th percentile of a frequency distribution of reference sites has been used to define the lower boundary of reference conditions (Figure 4.3). If the indicator (e.g. percent of all individuals that are in pollution-tolerant species or diatom-inferred total phosphorus concentration) is lowest in reference conditions, the 25th percentile of a frequency distribution of all sites or the 75th percentile of a frequency distribution of reference sites has been used
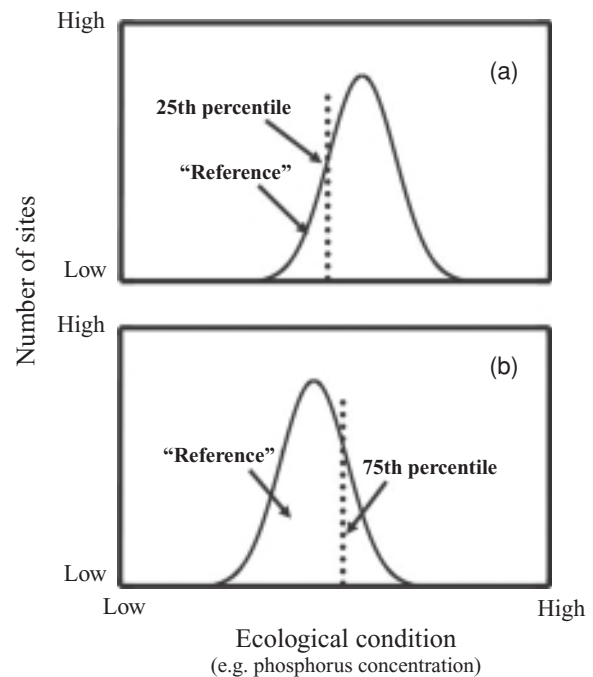

Figure 4.3 The frequency-distribution approach for establishing environmental criteria (Barbour *et al.*, 1999). The number of sites with different levels of an ecological attribute are plotted. In this case, the ecological attribute is a negative attribute, such as algal biomass or a pollutant; therefore sites with low levels of the ecological attribute characterize reference condition. Often, quartiles of frequency distributions are used to establish management targets. (A) The lower quartiles of frequency distributions that include all sites are assumed to provide reasonable restoration targets in a region in which substantial alteration of the landscape has occurred, but these quartiles may be overly protective in regions in which little alteration of the landscape has occurred. The upper quartile of frequency distributions using reference sites allows for some variation in measurement and actual expected value of expected condition; it is assumed to provide a balance of type I and II statistical errors for protecting the mean or median of reference conditions.

to define the upper-boundary of reference conditions. Several problems have been noted using the frequency-distribution method for establishing criteria for reference condition, such as arbitrary selection of percentiles and lack of established relationships between measures of ecological condition and human activities. Predictive models and non-linear relationships help resolve the latter problems.

Predictive models are used to define reference condition by relating measures of ecological condition to indicators of human alterations of watersheds that produce contamination or habitat alterations (Figure 4.4). For example, the minimally disturbed reference condition of TP can be predicted by relating TP concentrations to the percentage of agricultural and urban land use in watersheds (Dodds and Oakes 2004; Stevenson *et al.*,
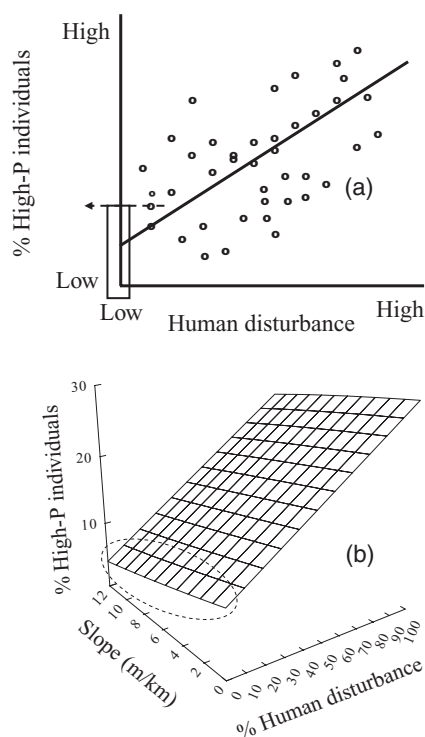
Figure 4.4 The predictive model approach (a & b) for characterizing reference condition. Using relationships between responses (e.g. percent high-P individuals) and indicators of human activities (e.g. percentages of urban and agricultural land use in watersheds) to determine natural or best attainable conditions. (a) Prediction of natural condition (y-intercept) using the regression relationship between biological condition and human activities assuming little natural variation in ecological condition. The arrow indicates the upper 75th percentile of the estimate of the y-intercept; i.e. when human disturbance is equal to zero. (b) Predictive model of natural condition (indicated by dashed oval) using a regression relationship between biological condition and human activities assuming considerable variation due to other natural factors (e.g. stream slope). Natural condition is calculated for a site with a model, measures of natural factors at assessed sites, and setting indicators of human activities to zero or an acceptably low value.

2008a). Similarly, the biologically inferred abiotic condition of minimally disturbed reference conditions can be determined by relating diatom indicators to the percentage of agricultural and urban land use in watersheds and accounting for effects of natural variation among streams on expected condition (Stevenson et al., 2009). The strength of this predictive modeling approach is that it enables modeling of natural, minimally disturbed, or even best attainable conditions without having to have a large number of reference sites, and it can account for natural variability among sites.

Reference condition can be explicitly or implicitly incorporated into indicators of ecological condition. In the EU, the WFD

calls for determination of the ecological quality ratio (EQR), which is the ratio between the observed status and expected status, where expected status is a minimally disturbed reference condition. Therefore the EQR explicitly measures deviation between the observed and expected condition. The EQR varies between 0 and 1 for low and high ecological quality, respectively. States in the USA commonly use indicators or multimetric indices of biological condition in characterizations of biological condition without explicitly incorporating the reference condition in the calculation; therefore interpreting the magnitude of deviation from reference condition requires an explicit statement of reference condition or only relative condition can be interpreted by comparison to other sites.

4.4.1.2 Ecological criteria and assessment There is an important distinction between reference condition and the management goals for waters, because the natural or minimally disturbed condition is not often a practical management goal in many landscapes, such as watersheds with extensive alteration by farming and urban activities. In the US CWA, biological integrity is the ultimate goal for US waters, but "the protection and propagation of fish, shellfish, and wildlife" is considered an interim goal, even though it falls short of biological integrity in most definitions used by states. This interim goal recognizes the need for a practical, but acceptable goal for protecting waters in the USA. The EU WFD calls for all surface waters to have "good ecological status," which is defined as "having biota consistent with only slight alterations from that expected in that absence of human impacts" (Kelly et al., 2008). Thus, in many cases, we need more than one level of ecological condition for setting appropriate management goals for the diversity of our waters (Davies and Jackson, 2006).

As a result, multiple management goals are needed for a more flexible and arguably more protective approach that enables different goals for different water bodies. In the USA, this approach has been referred to as tiered aquatic life "uses" (Davies and Jackson, 2006), and in the EU this is manifested in the different ecological-status categories (high, good, moderate, poor, and bad) and acceptability of the two higher goals (e.g. Kelly et al., 2008). The lower bounds of indicator values for each category of ecological condition become the targets for restoration and protection. These lower bounds are water-quality criteria in the USA, and they have similar meaning in the EU. Here we will refer to them as ecological criteria.

Several approaches have been used to establish ecological criteria. When there is just one acceptable use and associated criterion, then the frequency-distribution approach is often

satisfactory. When there is more than one level, governments have chosen simply to divide the range of conditions into a pre-specified number of categories. However, a more scientifically defensible approach is to relate loss of ecological condition to stressors and human activities, or what has been referred to as the pressure gradient. The challenge with this approach is to decide what specific levels of ecological condition should be chosen as criteria (or the boundaries of the ecological-status categories), especially when responses are linear.

Non-linear relationships between valued ecological attributes and stressors or measures of human activities are particularly valuable for establishing criteria for reference conditions, especially if they demonstrate thresholds (Muradian, 2001; Stevenson *et al.*, 2004a, 2008a; Figure 4.1). Thresholds help justify where to establish both stressor and biological criteria. With this approach, we first identify an indicator of a valued ecological attribute that non-linearly responds to increasing stressor levels, and then set a stressor criterion at a level of stressor that provides a margin of safety for protection of the valued attribute. Then we identify an indicator of valued attributes that responds linearly to increasing stressor levels, because it provides a sensitive and consistent response to changing stressor levels over the entire range of stressor conditions. Thus, a linear indicator provides an early warning of risk to a valued attribute that has a threshold response as stressor levels increase. Finally, the biological criterion for delineating reference condition can be established at a level that corresponds to the stressor criterion. If assessed biological condition decreases below the biological criterion, then the risk of unacceptable degradation of valued attributes is too high.

Other analytical approaches could be used to establish biological criteria or boundaries for defining ecological status. Kelly *et al.* (2008) relate the relative abundance of nutrient-sensitive and nutrient-tolerant taxa to an EQR, based on a trophic diatom index. They use the point where nutrient tolerant-species become more abundant than nutrient-sensitive species as a benchmark for the boundary between moderate and good ecological status. Kelly *et al.* argue that the benchmark for good status represents a point below which there is a shift in periphyton functioning; therefore it represents an objective and defensible criterion. Variations on this approach could evaluate metric thresholds along MMI-defined gradients, in which the MMI is a proxy for the human-disturbance gradient. The values of MMIs where thresholds in various metrics occur could be used as benchmarks for criteria for the tiered (successively higher) ecological-status categories (Figure 4.1).

The above methods for assessing ecological condition provide a set of diatom indicators that could be used complementarily in environmental assessments. Diatom indicators of valued ecological attributes with non-linear responses to stressors can be used to establish stressor criteria and ecological-status boundaries. Diatom indicators of biological condition with linear responses can be used to assess this condition with higher precision than indicators with non-linear responses. And finally, diatom indicators inferring stressor conditions could be used to complement measurement of stressors, to provide more precise and potentially accurate assessments of stressor conditions than simply measuring stressor conditions (e.g. water chemistry) on one sampling day (Stevenson 2006, unpublished data). Because of the high spatial and temporal variability in nutrient concentrations in streams and rivers, diatom indicators inferring nutrients could be very valuable as nutrient criteria in a multimetric index that includes actual measurements of nutrient concentrations.

4.4.2 Refining assessments by accounting for natural variability

Due to the great natural variability in ecological conditions of minimally disturbed streams and rivers, refining definitions of reference condition can increase precision in assessments (Figure 4.5). For example, streams in two geologically different settings could naturally have different pH and conductivity levels and therefore, naturally different diatom species occurring in them. Accounting for natural differences among ecosystems provides a more precise definition of reference condition, and consequently a more precise assessment of deviation from reference condition (Hawkins *et al.* 2000a). This concept can be integrated into all three assessment approaches described above.

Over the last 10 years, many conceptual and statistical approaches have been developed and evaluated for refining definitions of reference condition and incorporating them into assessments. Some of the earliest approaches for refining definitions of reference condition were *a-priori* separations of sites into groups by regions and stream size. Thus, before data were collected and analyzed, we assumed *a priori* that one or more naturally occurring ecological factors could be used to group sites. Regional variation in determinants of stream condition underpin the ecoregion approach, in which we assume that regional variation in geologic, climatic, and biogeographic factors explain significant amounts of variation in causal pathways and that they regulate diatom species composition (Biggs, 1995; Stevenson, 1997; Biggs *et al.*, 1998; Soininen, 2007). The
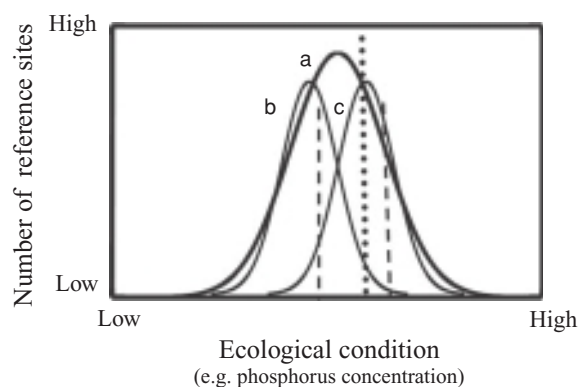
Figure 4.5 Refined frequency-distribution approach for defining reference condition. Separate distributions are developed for two or more classes of sites (e.g. b and c), where classes are based on natural variation in ecological condition or regional differences in management goals that may be based on the extent of human activities in watersheds. The vertical lines indicate the 75th percentiles of reference condition, with the longer dotted line for the aggregated distribution a and shorter left and right dashed lines for distributions b and c, respectively. Different criteria are established for each class of sites. The refined definition of reference condition based on natural factors provides more appropriate protection for different classes of sites, such that a higher level of protection is possible for one class of sites without being overly protective for the other class of sites.

ecoregion approach for refining the definition of reference condition has been used extensively in the USA for stream assessments. The approach has been criticized because relatively little variation in diatom species composition is explained by ecoregions (Pan *et al.* 2000) compared to the many variables which can vary naturally within ecoregions, such as stream pH, size, slope, substratum, watershed geology, and wetlands proportions of watersheds (Lavoie *et al.*, 2006; Cao *et al.*, 2007). However, Fore and Grafe (2002) successfully aggregated ecoregions to develop different metric expectations in different stream typologies based on altitude and regional land use.

More recently, refined characterizations of reference condition have been determined by using so-called *post-priori* approaches. These approaches are called *post priori* because after data are collected to develop indicators, they are analyzed to develop models for refining definitions of reference condition. After these indicators and models are developed, they are applied in assessments. The analytical methods vary greatly among all these approaches, which present a challenge to group and describe. One way to group methods is by establishing whether the methods call for separating reference sites from all sites in developing the model, which can limit usefulness when the numbers of sites in regions are small.

Kelly *et al.* (2008) developed site-specific expectations for indicator values by using regression analysis to determine effects of natural factors on Trophic Diatom Index (TDI) values across the United Kingdom. In this case, stream alkalinity was assumed to be independent of human activities in watersheds and the alkalinity explained the greatest variation in TDI among sites. The biological condition at a site can then be determined by comparing the observed TDI with the the expected TDI based on measured alkalinity at a site and the alkalinity model for TDI (Kelly *et al.*, 2008). Stevenson *et al.* (2009) used multiple regression to predict natural values of three trophic indices in Michigan streams, in which both total land use (percent agriculture and urban land use) and either wetland or geological conditions in watersheds were important. Natural factors explained almost as much variation in the trophic indices as total land use. Again, site-specific assessments were calculated as the difference between the observed and predicted trophic index values at a site, based on the wetland and geological conditions in watersheds and with land use set to zero. Cao *et al.* (2007) also developed an approach using site-specific predictions of the reference condition, but they used classification and regression tree (CART) and only reference sites to identify natural factors regulating metric values and to develop the predictive model of metrics in the reference condition.

As with the Cao *et al.* (2007) example, we can refine our expectations for species composition and metric values at a site by developing a better understanding of the natural factors that explain variation in diatom species composition using only reference sites. These models are like those used in the River Invertebrate Prediction and Classification System (RIVPACS, *sensu* Moss *et al.* 1987; Hawkins *et al.* 2000b) and they use cluster analysis to identify groups of sites with similar species composition; then discriminant function analysis is used to identify distinguishing natural factors among the groups of sites and to build models for assigning sites with a predicted condition (Chessman *et al.*, 1999; Cao *et al.*, 2007; Carlisle *et al.*, 2008). The RIVPACS models usually assess the observed number of species typical of reference sites compared to the expected number at a site with similar typology. The value of RIVPACS models in the assessment of the proportion of reference taxa at an assessed site, which is a direct measure of the biodiversity, a valued element of biological condition and it could be applied consistently in all assessments (Hawkins *et al.*, 2000b); however the hump-shaped relationships between species numbers in counts and environmental gradients can cause problems with this metric (Stevenson *et al.*, 2008a). A similar approach is being

developed using self-organizing maps to group reference sites with species composition; then discriminant function analysis and similar methods (such as multilayer perception) can be used to predict what an assemblage should be in the absence of anthropogenic impacts (Tison *et al.*, 2007). In a very simple version of this approach, Grenier *et al.* (2006) found that the natural conditions of sites in the St. Lawrence basin of Quebec, Canada, were best explained by differences in pH; then Lavoie *et al.* (2006) developed separate metric expectations in streams with low and high pH.

4.4.3 Stressors condition and diagnosis

Many diatom indicators that have been developed in the last century really infer the stressor conditions in a habitat, rather than assess the biological condition (Stevenson 2006). Thus, diatom indicators of stressor condition can be used to complement actual measures of stressor condition. This has been particularly important in paleoecology and such efforts as inferring pH, trophic status, and climate change in lakes for which diatoms are especially valuable proxies of past conditions that can not be measured directly (Charles *et al.* 1990, Dixit *et al.* 1999, Smol *et al.* 2005, and several chapters in this volume). Diatom indicators are also valuable in streams and rivers because of the great temporal variability in some variables, such as nutrient concentrations. Stevenson (2006) summarized results of unpublished data showing how diatom indicators of TP concentration can be more precise and accurate indicators of TP availability than measured TP concentrations (Figure 4.2).

Diatom indicators of nutrient concentrations, dissolved oxygen, organic matter, pH, conductivity, and sediments have been developed for streams (Table 4.1). The diatom indicators for nutrients, dissolved oxygen, and pH should be particularly useful because these factors vary so much on diurnal and weekly timescales due to weather and diurnal variation in metabolic processes. However, diatom indicators of stressor conditions should be rigorously tested to confirm their causal linkage to the stressor that they are designed to indicate. Stevenson *et al.* (2008b) illustrated this problem showing the challenge of identifying causal versus correlational relationships for diatom indicators of stressors in large regional surveys. Despite reasonable levels of precision in indicators of pH, conductivity, fine sediments, and embeddedness, these diatom indicators were actually more precisely related to percent disturbance in watersheds than the stressors for which the weighted-average models were developed. Covariation among environmental factors can be a significant problem for indicator development in large regional projects. Subdividing large data sets may help reduce covariation and the development of more causally related indicators.

Diatom indicators of stressors can be used in stressor diagnosis in a number of ways. First, diatom indicators of stressors could be used to refine stressor–response relationships, such as responses of algal biomass to TP (Stevenson 2006). By refining stressor–response relationships, they can help resolve threshold relationships and establish expected conditions or criteria for stressors (Figure 4.2). Second, they could be used for the development of criteria for expected stressor conditions. In addition to helping establish the specific level of a stressor that should be designated as the water-quality criterion, the inferred stressor condition by the diatom indicator, as well as the analytically measured stressor condition, could be used as water-quality criteria. For example, we could use a diatom indicator of TP concentration as well as measured TP concentrations to determine whether TP criteria were being violated or not. Using multiple lines of evidence, or multiple indicators, usually increases the precision and accuracy of assessments. Finally, we need to assess stressor conditions in the habitat for stressor diagnosis. Use of the diatom-inferred stressor condition alone or in combination as a multimetric index with an actual measured condition could improve the accuracy and precision of stressor assessments.

Stressor diagnosis is critical for both protection and restoration of ecosystems (Stevenson *et al.*, 2004b). Beyers (1998) amended and refined the postulates of Hill (1965) to list a set of criteria that should be considered when trying to infer the cause of an un-replicated environmental impact. These criteria are: extent of alterations (strength), observation by other investigators (consistency), unique effects specific to the stressor (specificity), exposure to the stressor prior to the observed effect (temporality), the relationship between the magnitude of the stressor and the effect (stressor–response relationship), plausibility of causal linkage, experimental evidence, transferability of stressor/effects elsewhere, causal hypothesis consistent with existing knowledge, and exposure sufficient for the effect. Thus, proper identification of the contaminant and habitat alterations that are causing problems can be facilitated with diatom indicators of stressors. First, if stressor levels are greater than stressor criteria, then those stressors should be targeted for remediation. Second, ratios and differences between stressor levels and stressor criteria can be used to rank the importance of different stressors. Ratios between stressor levels and criteria follow the concept of toxic units in the toxicological literature (Stevenson *et al.*, 2004b). Differences between stressor levels and criteria have been related

to the sustainability and restorability of habitats (Stevenson, 1998). Third, and more recently, the probability of losing valued attributes has been related to stressor levels in a risk-based approach (Paul and McDonald, 2005; Rollins and Stevenson, unpublished manuscript). Thus, diatom indicators of stressors can be used in stressor diagnosis in the same way as actual measurements of the physical, chemical, and biological alterations of habitats by humans.

4.5 Concluding remarks

Developing approaches and indices for environmental assessment is an interactive process between scientists and policy makers. Interactions should focus on furthering scientists' understanding of policy issues and environmental problems and on helping policy makers translate their goals into testable hypotheses and practical approaches for environmental assessment and problem solving. Priorities of some policy makers have been directed toward understanding relationships between land use, physical and chemical changes in streams and rivers, and ecological responses. Land-use planning and zoning are important strategies for slowing environmental degradation in areas under development pressure from urban and suburban sprawl. Assessing watershed-scale changes in stream and river conditions could be valuable for inferring the land-use effects and the geological and climatic factors that make watersheds sensitive or tolerant to land-use changes (Robinson et al., 1995; Richards et al., 1996; Kutka and Richards, 1996).

Ecological theory should become a more important foundation for the environmental assessments and the indices used in environmental assessments. For example, Ruth Patrick's pioneering work on species diversity as an indicator of water quality was well founded in the ecological theory that was being explored at the time (Patrick et al., 1954; Patrick and Strawbridge, 1963). By placing research in an ecological context and testing a broader ecological theory, such as Odum's predictions for stressed ecosystems (Odum et al., 1979; Odum, 1985), as well as specific diatom-based hypotheses, results of our research become more transferable to assessments with other organisms and to assessments of other habitats with diatoms.

Diatoms are valuable indicators of biological condition and the environmental factors that impair rivers and streams. Public concern often focuses on the biodiversity of other organisms, but partly because they do not appreciate the diversity, beauty, and ecology of algae, particularly diatoms. Greater efforts should be made to inform the public and develop their appreciation for diatoms in aquatic ecosystems so they know

what valued attributes are at risk. In addition, development of diatom indicators of ecosystem function and services provides a direct linkage to valued ecological attributes. Relating diatom assessments to questions being asked by stakeholders or to questions that they should be asking is essential for sustainable diatom monitoring programs. Great investments are being made in diatom assessment programs, on the scale of millions of dollars per year. They require our accountability and creativity to meet the needs of the stakeholders.

4.6 Summary

Diatoms have a long history of use in assessing the ecological integrity of streams. Diatom assemblages respond rapidly and sensitively to environmental change and provide highly informative assessments of the biotic integrity of streams and rivers and causes of ecosystem impairment. Periphytic diatoms from natural and artificial substrata are usually sampled from streams and small rivers, but plankton provides valuable assessments of conditions in large rivers. Structural and functional characteristics of diatom communities can be used in bioassessments, but relative abundances of diatom genera and species are usually used as the most valuable characteristics of diatom assemblages for bioassessment. Using these characteristics, multimetric indices of biotic and ecological condition have been developed that enable use of diatom assemblages in risk assessment and management of stream and river ecosystems.

Exciting new approaches for indicator refinements and applications provide opportunities for research at the cutting edge of applied ecological science. With near-natural condition as a management target for many governments, the refining of indicators to account for natural variability and the differing response to stressors in different types of streams and rivers present challenges for future study. The refining of indicators is also dependent upon sound analytical technique, including better knowledge of diatom taxonomy and communication of that taxonomy. Understanding threshold responses in diatom assemblages along environmental gradients is valuable for justifying specific management targets. Thus, a more complete understanding of the taxonomy and ecology of diatoms remains at the foundation of sound science for environmental assessments and advancements in that science.

colleagues and students for the many rigorous and earnest debates about algal ecology and ecological assessment. In particular, we thank Martyn Kelly for his insightful contributions while reviewing this chapter. Research enabling this synthesis was supported by grants and contracts from the US Environmental Protection Agency. We dedicate this chapter to Dr. Charles Reimer for his inspiration and teaching.

References

Alatalo, R. V. (1981). Problems in the measurement of evenness in ecology. *Oikos*, **37**, 199–204.

Aloi, J. E. (1990). A critical review of recent freshwater periphyton methods. *Canadian Journal of Fisheries and Aquatic Sciences*, **47**, 656–70.

Alverson, A. J., Manoylov, K. M., and Stevenson, R. J. (2003). Laboratory sources of error for algal community attributes during sample preparation and counting. *Journal of Applied Phycology*, **15**, 357–69.

American Public Health Association (APHA) (1998). *Standard Methods for the Evaluation of Water and Wastewater*, 20th edition, American Public Health Association, Washington, DC.

Amoros, C. and Van Urk, G. (1989). Palaeoecological analyses of large rivers: some principles and methods. In *Historical Change of Large Alluvial Rivers: Western Europe*, ed. G. E. Petts, H. Möller and A. L. Roux, Chichester: John Wiley & Sons, pp. 143–65.

Andrén, C. & Jarlman, A. (2008). Benthic diatoms as indicators of acidity in streams. *Fundamental and Applied Limnology*, **173**, 237–53.

Archibald, R. E. M. (1972). Diversity in some South African diatom assemblages and its relation to water quality. *Water Research*, **6**, 1229–38.

Asai, K. (1996). Statistical classification of epilithic diatom species into three ecological groups relating to organic water pollution. (1) Method with coexistence index. *Diatom*, **10**, 13–34.

Bahls, L. L. (1993). Periphyton Bioassessment Methods for Montana Streams. Water Quality Bureau, Department of Health and Environmental Sciences, Helena, MT.

Barbour, M. T., Gerritsen, J., Snyder, B. D., and Stribling, J. B. (1999). Revision to Rapid Bioassessment Protocols for Use in Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. EPA 841-D-97–002, Washington, DC: United States Environmental Protection Agency.

Beaver, J. (1981). Apparent Ecological Characteristics of Some Common Freshwater Diatoms. Rexdale, Ontario: Ontario Ministry of the Environment.

Bennion, H., Juggins, S., and Anderson, N. J. (1996). Predicting epilimnetic phosphorus concentrations using an improved diatom-based transfer function and its application to lake eutrophication management. *Environmental Science & Technology*, **30**, 2004–7.

Besse-Lototskaya, A., Verdonschot, P. F. M., and Sinkeldam, J. A. (2006). Uncertainty in diatom assessment: sampling, identification and counting variation. *Hydrobiologia*, **566**, 247–60.

Beyers, D. W. (1998). Causal inference in environmental impact studies. *Journal of the North American Benthological Society*, **17**, 367–73.

Biggs, B. J. F. (1990). Use of relative specific growth rates of periphytic diatoms to assess enrichment of a stream. *New Zealand Journal of Marine and Freshwater Research*, **24**, 9–18.

Biggs, B. J. F. (1995). The contribution of flood disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. *Freshwater Biology*, **33**, 419–38.

Biggs, B. J. F. (1996). Patterns of benthic algae in streams. In *Algal Ecology: Freshwater Benthic Ecosystems*, ed. R. J. Stevenson, M. Bothwell, and R. L. Lowe, San Diego, CA: Academic Press, pp. 31–55.

Biggs, B. J. F. (2000). Eutrophication of streams and rivers: dissolved nutrient- chlorophyll relationships for benthic algae. *Journal of the North American Benthological Society*, **19**, 17–31.

Biggs, B. J. F., Stevenson, R. J. & Lowe, R. L. (1998). A habitat matrix conceptual model for stream periphyton. *Archiv für Hydrobiologie*, **143**, 21–56.

Blanck, H. (1985). A simple, community level, ecotoxicological test system using samples of periphyton. *Hydrobiologia*, **124**, 251–61.

Bothwell, M. L. (1989). Phosphorus-limited growth dynamics of lotic periphytic diatom communities: areal biomass and cellular growth rate responses. *Canadian Journal of Fisheries and Aquatic Sciences*, **46**, 1293–301.

Bothwell, M. L., Sherbot, D., Roberge, A. C., and Daley, R. J. (1993). Influence of natural ultraviolet radiation on lotic periphytic diatom community growth, biomass accrual, and species composition: short-term versus long-term effects. *Journal of Phycology*, **29**, 24–35.

Brugam, R., McKeever, K., and Kolesa, L. (1998). A diatom-inferred water depth reconstruction for Upper Peninsula, Michigan lake. *Journal of Paleolimnology*, **20**, 267–76.

Brugam, R. & Patterson, C. (1983). The A/C (Araphidineae/Centrales) ratio in high and low alkalinity lakes in eastern Minnesota. *Freshwater Biology*, **13**, 47–55.

Butcher, R. W. (1947). Studies in the ecology of rivers. IV. The algae of organically enriched water. *Journal of Ecology*, **35**, 186–91.

Cairns, J. Jr. and Kaesler, R. L. (1969). Cluster analysis of Potomac River survey stations based on protozoan presence–absence data. *Hydrobiologia*, **34**, 414–32.

Cao, Y., Hawkins, C. P., Olson, J., and Kosterman, M. A. (2007). Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. *Journal of North American Benthological Society*, **26**, 566–84.

Carlisle, D. M., Hawkins, C. P., Meador, M. R., Potapova, M., and Falcone, J. (2008). Biological assessments of Appalachian streams based on predictive models for fish, macroinvertebrate, and diatom assemblages. *Journal of North American Benthological Society*, **27**, 16–37.

Cattaneo, A. & Amireault, M. C. (1992). How artificial are artificial substrata for periphyton? *Journal of the North American Benthological Society*, **11**, 244–56.

CEN (Comité Européen de Normalisation) (2003). Water quality – guidance standard for the routine sampling and pretreatment

of benthic diatoms from rivers. European Standard EN 13946, Brussels.

CEN (2004). Water quality – guidance standard for the identification, enumeration and interpretation of benthic diatom samples from running waters. European Standard EN 14407.

Charles, D. F. (1985). Relationships between surface sediment diatom assemblages and lake-water characteristics in Adirondack lakes. *Ecology*, **66**, 994–1011.

Charles, D. F., Binford, M. W., Furlong, E. T., *et al.* (1990). Paleoecological investigation of recent lake acidification in the Adirondack Mountains, N. Y. *Journal of Paleolimnology*, **3**, 195–241.

Charles, D. F., Knowles, C., and Davis, R.S. (eds.) (2002). Protocols for the analysis of algal samples collected as part of the U.S. Geological Survey National Water-Quality Assessment Program. Report No. 02–06, Patrick Center for Environmental Research, The Academy of Natural Sciences, Philadelphia, PA. See http://diatom.acnatsci.org/nawqa/pdfs/ProtocolPublication.pdf

Chessman, B. C. (1986). Diatom flora of an Australian river system: spatial patterns and environmental relationships. *Freshwater Biology*, **16**, 805–19.

Chessman, B., Growns, I., Currey, J., and Plunkett-Cole, N. (1999). Predicting diatom communities at the genus level for the rapid biological assessment of rivers. *Freshwater Biology*, **41**, 317–31.

Chessman, B. C., Bate, N., Gell, P. A., and Newall, P. (2007). A diatom species index for bioassessment of Australian rivers. *Marine and Freshwater Research*, **58**, 542–57.

Cholnoky, B. J. (1968). *Okologie der Diatomeen in Binnengewässern*. Lehre: Cramer.

Clarke, K.R. & Gorley, R.N. (2006). PRIMER v6: Users Manual/Tutorial. PRIMER-E, Plymouth.

Coste, M. (1982). *Etude des méthodes biologiques d'appréciation quantitative de la qualité des eaux*. Lyon: CEMAGREF Division Qualité des Eaux, Agence de L'eau Rhône-Méditerrané Corse.

Coste, M., Bosca, C., and Dauta, A. (1991). Use of algae for monitoring rivers in France. In *Use of Algae for Monitoring Rivers*, ed. B. A. Whitton, E. Rott, and G. Friedrich, Innsbruck: Universität Innsbruck, pp. 75–88.

Cremer, H., Buijse, A., Lotter, A., Oosterberg, W., and Staras, M. (2004). The palaeolimnological potential of diatom assemblages in floodplain lakes of the Danube Delta, Romania: a pilot study. *Hydrobiologia*, **513**, 7–26.

Crumpton, W. G. (1987). A simple and reliable method for making permanent mounts of phytoplankton for light and fluorescence microscopy. *Limnology and Oceanography*, **32**, 1154–9.

Davies, S. P. & Jackson, S. K. (2006). The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. *Ecological Applications*, **16**, 1251–66.

Descy, J. P. (1979). A new approach to water quality estimation using diatoms. *Nova Hedwigia*, **64**, 305–23.

Descy, J. P. & Mouvet, C. (1984). Impact of the Tihange nuclear power plant on the periphyton and the phytoplankton of the Meuse River (Belgium). *Hydrobiologia*, **119**, 119–28.

DeYoe, H. R., Lowe, R. L., and Marks, J. C. (1992). The effect of nitrogen and phosphorus on the endosymbiot load of *Rhopalodia gibba* and *Epithemia turgida* (Bacillariophyceae). *Journal of Phycology*, **23**, 773–7.

Dixit, S. S., Smol, J. P., Charles, D. F., *et al.* (1999). Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms. *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 131–52.

Dodds, W. K. & Oakes, R. M. (2004). A technique for establishing reference nutrient concentrations across watersheds affected by humans. *Limnology and Oceanography: Methods*, **2**, 331–41.

Dodds, W. K., Smith, V. H., and Zander, B. (1997). Developing nutrient targets to control benthic chlorophyll levels in streams: a case study of the Clark Fork River. *Water Research*, **31**, 1738–50.

Dufrêne, M. & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*, **67**, 345–66.

Ector, L., Kingston, J. C., Charles, D. F., *et al.* (2004). Workshop report: freshwater diatoms and their role as ecological indicators. In *Proceedings of the Seventeenth International Diatom Symposium 2002*, ed. M. Poulin, Bristol: Biopress Ltd, pp. 469–80.

European Union (EU) (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. The European Parliament and the Council of the European Union. *Official Journal of the European Communities*, **L 327/1**, 1–72.

Fabri, R. & Leclercq, L. (1984). *Etude écologique des riviéres du nord du massif Ardennais (Belgique): flore et végétation de diatomées et physico-chimie des eaux*. 1. Robertville: Station scientifique des Hautes Fagnes.

Falasco, E., Bona, F., Badino, G., Hoffmann, L., and Ector, L. (2009). Diatom teratological forms and environmental alterations: a review. *Hydrobiologia*, **623**, 1–35

Fjerdingstad, E. (1950). The microflora of the River Molleaa with special reference to the relation of benthic algae to pollution. *Folia Limnologica Scandanavica*, **5**, 1–123.

Fore, L. & Grafe, C. (2002). Using diatoms to assess the biological condition of large rivers in Idaho (U.S.A.). *Freshwater Biology*, **47**, 2015–37.

Fore, L. S., Karr, J. R., and Conquest, L. L. (1994). Statistical properties of an index of biotic integrity used to evaluate water resources. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 1077–87.

Gale, W. F., Gurzynski, A. J., and Lowe, R. L. (1979). Colonization and standing crops of epilithic algae in the Susquehanna River, Pennsylvania. *Journal of Phycology*, **15**, 117–23.

Gell, P., Bulpin, S., Wallbrink, P., Hancock, G., and Bickford, S. (2005). Tareena Billabong – a palaeolimnological history of an ever-changing wetland, Chowilla Floodplain, lower Murray-Darling Basin, Australia. *Marine and Freshwater Research*, **56**, 441–56.

Genter, R. B. (1996). Ecotoxicology of inorganic chemical stress to algae. In *Algal Ecology: Freshwater Benthic Ecosystems*. ed. R. J. Stevenson, M. Bothwell and R. L. Lowe, San Diego, CA: Academic Press, pp. 403–68.

Gillett, N., Pan, Y. & Parker, C. (2009). Should only live diatoms be used in the bioassessment of small mountain streams? *Hydrobiologia*, **620**, 135–47.

Gómez, N. & Licursi, M. (2001). The Pampean Diatom Index (DPI) for assessment of rivers and streams in Argentina. *Aquatic Ecology*, **35**, 173–81.

Grenier, M., Campeau, S., Lavoie, I., Park, Y.-S. & Lek, S. (2006). Diatom reference communities in Québec (Canada) streams based on Kohonen self-organized maps and multivariate analyses. *Canadian Journal of Fisheries and Aquatic Sciences*, **63**, 2087–106.

Griffith, M. B., Hill, B. H., McCormick, F. H., *et al.* (2005). Comparative application of indices of biotic integrity based on periphyton, macroinvertebrates, and fish to southern Rocky Mountain streams. *Ecological Indicators*, **5**, 117–36.

Hawkins, C. P., Norris, R. H., Gerritsen, J., *et al.* (2000a). Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society*, **19**, 541–56.

Hawkins, C. P., Norris, R. H., Hogue, J. N. & Feminella, J. W. (2000b). Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications*, **10**(5), 1456–77.

Healey, F. P. and Henzel, L. L. (1979). Fluorometric measurement of alkaline phosphatase activity in algae. *Freshwater Biology*, **9**, 429–39.

Hering, D., Johnson, R. K., Kramm, S., *et al.* (2006). Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response due to stress. *Freshwater Biology*, **51**, 1757–85.

Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, **58**, 295–300.

Hill, B. H., Herlihy, A. T., Kaufmann, P. R. & Sinsabaugh, R. L. (1998). Sediment microbial respiration in a synoptic survey of mid-Atlantic region streams. *Freshwater Biology*, **39**, 493–501.

Hill, B., Herlihy, A., Kaufmann, P., *et al.* (2000). Use of periphyton assemblage data as an index of biotic integrity. *Journal of North American Benthological Society*, **19**, 50–67.

Hill, B. H., Lazorchak, J. M., McCormick, F. H., and Willingham, W. T. (1997). The effects of elevated metals on benthic community metabolism in a Rocky Mountain stream. *Environmental Pollution*, **95**, 183–90.

Hill, B. H., Stevenson, R. J., Pan, Y., *et al.* (2001). Correlations of stream diatoms with their environment: a comparison of genus-level and species-level identifications. *Journal of the North American Benthological Society*, **20**, 299–310.

Hill, M. O. (1979). TWINSPAN-A FORTRAN Program for Detrended Correspondence Analysis and Reciprocal Averaging. Cornell University, Ithaca, NY.

Hill, W. R. & Boston, H. L. (1991). Community development alters photosynthesis-irradiance relations in stream periphyton. *Limnology and Oceanography*, **36**, 1375–89.

Hilsenhoff, W. L. (1988). Rapid field assessment of organic pollution with a family level biotic index. *Journal of the North American Benthological Society*, **7**, 65–8.

Hoagland, K. D., Carder, J. P. & Spawn, R. L. (1996). Effects of organic toxic substances. In *Algal Ecology: Freshwater Benthic Ecosystems*, ed. R. J. Stevenson, M. Bothwell and R. L. Lowe, San Diego, CA, Academic Press, pp. 469–97.

Humphrey, K. P. & Stevenson, R. J. (1992). Responses of benthic algae to pulses in current and nutrients during simulations of sub-scouring spates. *Journal of the North American Benthological Society*, **11**, 37–48.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, **52**, 577–86.

Hurlbert, S. J. (1984). Pseudoreplication and design of ecological field experiments. *Ecological Monographs*, **54**, 187–211.

Hustedt, F. (1957). Die Diatomeenflora des Flusssystems der Weser im Gebiet der Hansestadt Bremen. *Bremen: Abhandlungen Naturwissenschaftlichen Verein*, **34**(3), 181–440.

Johnson, R. K., Hering, D., Furse, M. T., and Clarke, R. T. (2006). Detection of ecological change using multiple organism groups: metrics and uncertainty. *Hydrobiologia*, **566**, 115–37.

Jüttner, I., Rothfritz, H., and Omerod, S. J. (1996). Diatoms as indicators of river water quality in the Nepalese Middle Hills with consideration of the effects of habitat-specific sampling. *Freshwater Biology*, **36**, 475–86.

Kahlert, M., Albert, R.-L., Anttila, E.-L., *et al.* (2009). Harmonization is more important than experience – results of the first Nordic–Baltic diatom intercalibration exercise 2007 (stream monitoring). *Journal of Applied Phycology*, **21**, 471–82.

Karr, J. R. (1981). Assessment of biotic integrity using fish communities. *Fisheries*, **6**, 21–7.

Karr, J. R. & Chu, E. W. (1997). Biological Assessment: Using Multimetric Indexes Effectively. United States Environmental Protection Agency, Washington, D. C.

Karr, J. R. & Dudley, D. R. (1981). Ecological perspective on water quality goals. *Environmental Management*, **5**, 55–68.

Kawecka, B. (1993). Ecological characteristics of sessile algal communities in streams flowing from the Tatra Mountains in the area of Zakopane (southern Poland) with special consideration of their requirements with regard to nutrients. *Acta Hydrobiologica*, **35**, 295–306.

KDOW (Kentucky Division of Water) (2008). Standard Methods for Assessing Biological Integrity of Surface Waters in Kentucky. Lexington, KY: Kentucky Division of Water.

Kelly, M., Bennett, C., Coste, M., *et al.* (2009a). A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. *Hydrobiologia*, **621**, 169–82.

Kelly, M., Bennion, H. Burgess, A., *et al.* (2009b). Uncertainty in ecological status assessments of lakes and rivers using diatoms. *Hydrobiologia*, **633**, 5–15.

Kelly, M., Juggins, S., Guthrie, R., *et al.* (2008). Assessment of ecological status in U.K. rivers using diatoms. *Freshwater Biology*, **53**, 403–22.

Kelly, M. & Whitton, B. A. (1995). The trophic diatom index: a new index for monitoring eutrophication in rivers. *Journal of Applied Ecology*, **7**, 433–44.

Kelly, M. G. (1998). Use of the trophic diatom index to monitor eutrophication in rivers. *Water Research*, **32**, 236–42.

Kelly, M. G. (2001). Use of similarity measures for quality control of benthic diatom samples. *Water Research*, **35**, 2784–88.

Kelly, M. G. (2003). Short term dynamics of diatoms in an upland stream and implications for monitoring eutrophication. *Environmental Pollution*, **125**, 117–22.

Kelly, M. G., Cazaubon, A., Coring, E., *et al.* (1998). Recommendations for the routine sampling of diatoms for water quality assessments in Europe. *Journal of Applied Phycology*, **10**, 215–24.

Kelly, M. G., Penny, C. J., and Whitton, B. A. (1995). Comparative performance of benthic diatom indices used to assess river water quality. *Hydrobiologia*, **302**, 179–88.

Kelly, M. G. & Whitton, B. A. (1989). Interspecific differences in Zn, Cd and Pb accumulation by freshwater algae and bryophytes. *Hydrobiologia*, **175**, 1–11.

Kim, B. K., Jackman, A. P., and Triska, R. J. (1990). Modeling transient storage and nitrate uptake kinetics in a flume containing a natural periphyton community. *Water Resources Research*, **26**, 505–15.

Kolkwitz, R. & Marsson, M. (1908). Ökologie der pflanzliche Saprobien. *Berichte der Deutsche Botanische Gesellschaften*, **26**, 505–19.

Kovács, C. M., Kahlert, M., and Padisák, J. (2006). Benthic diatom communities along pH and TP gradients in Hungarian and Swedish streams. *Journal of Applied Phycology*, **18**, 105–17.

Kutka, F. J. & Richards, C. (1996). Relating diatom assemblage structure to stream habitat. *Journal of the North American Benthological Society*, **15**, 469–80.

Lamberti, G. A. (1996). The role of periphyton in benthic food webs. In *Algal Ecology: Freshwater Benthic Ecosystems*, ed. R. J. Stevenson, M. Bothwell and R. L. Lowe, San Diego, CA: Academic Press, pp. 533–72.

Lange-Bertalot, H. (1979). Pollution tolerance of diatoms as a criterion for water quality estimation. *Nova Hedwigia* **64**, 285–304.

Lavoie, I., Campeau, S., Grenier, M., and Dillon, P. J. (2006). A diatom-based index for the biological assessment of eastern Canadian rivers: an application of correspondence analysis (CA). *Canadian Journal of Fisheries and Aquatic Sciences*, **8**, 1793–811.

Lavoie, I., Hamilton, P. B., Wang, Y. K., Dillon, P. J., and Campeau, S. (2009). A comparison of stream bioassessment in Québec (Canada) using six European and North American diatom-based indices. *Nova Hedwigia*, **135**, 37–56.

Lazorchak, J. M., Klemm, D. J., and Peck, D. V. (1998). Environmental Monitoring and Assessment Program – Surface Waters: field operations and methods for measuring the ecological condition of wadeable streams. EPA-620-R-94–004F, Washington, DC: United States Environmental Protection Agency.

Leland, H. V. (1995). Distribution of phytobenthos in the Yakima River basin, Washington, in relation to geology, land use and other environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 1108–29.

Lenat, D. R. (1993). A biotic index for the southeastern United States: derivation and list of tolerance values, with criteria for assigning water quality ratings. *Journal of the North American Benthological Society*, **12**, 279–90.

Lobo, E. A., Callegaro, V. L. M., Hermany, G., Gómez, N., and Ector, L. (2004). Review of the use of microalgae in South America for monitoring rivers, with special reference to diatoms. *Vie Milieu*, **54**, 105–14.

Lowe, R. L. (1974). Environmental Requirements and Pollution Tolerance of Freshwater Diatoms. US Environmental Protection Agency, EPA-670/4–74-005, Cincinnati, OH.

Lowe, R. L. & Pan, Y. (1996). Benthic algal communities and biological monitors. In *Algal Ecology: Freshwater Benthic Ecosystems*, ed. R. J. Stevenson, M. Bothwell and R. L. Lowe, San Diego, CA: Academic Press, pp. 705–39.

Manoylov, K. M. & Stevenson, R. J. (2006). Density-dependent algal growth along N and P nutrient gradients in artificial streams. In *Advances in Phycological Studies*, ed. N. Ognjanova-Rumenova and K. Manoylov, Moscow: Pensoft Publishers, pp. 333–52.

Mayer, M. S. and Likens, G. E. (1987). The importance of algae in a shaded headwater stream as food for an abundant caddisfly (Trichoptera). *Journal of the North American Benthological Society*, **6**, 262–9.

McCormick, P. V. & Stevenson, R. J. (1989). Effects of snail grazing on benthic algal community structure in different nutrient environments. *Journal of the North American Benthological Society*, **82**, 162–72.

McCune, B., and Grace, J. B. (2002). Analysis of ecological communities. MjM software design, Gleneden Beach, OR.

McFarland, B. H., Hill, B. H., and Willingham, W. T. (1997). Abnormal *Fragilaria* spp. (Bacillariophyceae) in streams impacted by mine drainage. *Journal of Freshwater Ecology*, **12**, 141–9.

Moore, W. W. and McIntire, C. D. (1977). Spatial and seasonal distribution of littoral diatoms in Yaquina Estuary, Oregon (U.S.A.). *Botanica Marina*, **20**, 99–109.

Moss, D., Furse, M. T., Wright, J. F., and Armitage, P. D. (1987). The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology*, **17**, 41–52.

Moulton, S. R., Kennen, J. G., Goldstein, R. M., and Hambrook, J. A. (2002). Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program. Open-file Report 02–150, United States Geological Survey, Reston, VA.

Mulholland, P. J. (1996). Role of nutrient cycling in streams. In *Algal Ecology: Freshwater Benthic Ecosystems*, ed. R. J. Stevenson, M. Bothwell, and R. L. Lowe, San Diego, CA: Academic Press, pp. 609–39.

Mulholland, P. J. & Rosemond, A. D. (1992). Periphyton response to longitudinal nutrient depletion in a woodland stream: evidence of upstream-downstream linkage. *Journal of the North American Benthological Society*, **11**, 405–19.

Müller-Haeckel, A. and Håkansson, H. 1978. The diatom-flora of a small stream near Abisko (Swedish Lapland) and its annual periodicity, judged by drift and colonization. *Archiv für Hydrobiologie*, **84**, 199–217.

Muradian, R. (2001). Ecological thresholds: a survey. *Ecological Economics*, **38**, 7–24.

Newbold, J. D., Elwood, J. W., O'Neill, R. V., and Van Winkle, W. (1981). Measuring nutrient spiralling in streams. *Canadian Journal of Fisheries and Aquatic Sciences*, **38**, 680–3.

Niederlehner, B. R. & Cairns, J. C., Jr. (1994). Consistency and sensitivity of community level endpoints n microcosm tests. *Journal of Aquatic Ecosystem Health*, **3**, 93–9.

Norris, R. H. and Norris, K. R. (1995). The need for biological assessment of water quality: Australian perspective. *Australian Journal of Ecology*, **20**, 1–6.

Odum, E. P. (1985). Trends expected in stressed ecosystems. *BioScience*, **35**, 412–22.

Odum, E. P., Finn, J. T., and Franz, E. H. (1979). Perturbation theory and the subsidy-stress gradient. *BioScience*, **29**, 349–52.

Oksanen, J. (2004). 'Vegan' Community Ecology Package: ordination methods and other functions for community and vegetation ecologists. University of Oulu, Oulu, Finland.

Palmer, C. M. (1962). *Algae in Water Supplies.*, Washington, DC: US Department of Health, Education and Welfare.

Palmer, C. M. (1969). A composite rating of algae tolerating organic pollution. *Journal of Phycology*, **5**, 78–82.

Pan, Y. D., Stevenson, R. J., Hill, B. H., Herlihy, A. T., and Collins, G. B. (1996). Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. *Journal of the North American Benthological Society*, **15**, 481–95.

Pan, Y. D., Stevenson, R. J., Hill, B. H., and Herlihy, A. T. (2000). Ecoregions and benthic diatom assemblages in Mid-Atlantic Highlands streams, USA. *Journal of the North American Benthological Society*, **19**, 518–40.

Pantle, R. & Buck, H. (1955). Die biologische Überwachung der Gewasser und die Darstellung der Ergebnisse. *Gas- und Wasserfach*, **96**, 604.

Pappas, J. L. & Stoermer, E. F. (1996). Quantitative methods for determining a representative algal count. *Journal of Phycology*, **32**, 693–6.

Passy, S. I. & Bode, R. W. (2004). Diatom model affinity (DMA), a new index for water quality assessment. *Hydrobiologia*, **524**, 241–51.

Patrick, R. (1949). A proposed biological measure of stream conditions based on a survey of the Conestoga Basin, Lancaster County, Pennsylvania. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **101**, 277–341.

Patrick, R. (1961). A study of the numbers and kinds of species found in rivers of the Eastern United States. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **113**, 215–58.

Patrick, R. (1973). Use of algae, especially diatoms, in the assessment of water quality. In *Biological Methods for the Assessment of Water Quality*, ASTM STP 528, Philadelphia, PA: American Society for Testing and Materials, pp. 76–95.

Patrick, R., Hohn, M. H., and Wallace, J. H. (1954). A new method for determining the pattern of the diatom flora. *Notulae Naturae*, No. 259.

Patrick, R. and Strawbridge, D. (1963). Variation in the structure of natural diatom communities. *The American Naturalist*, **97**, 51–7.

Paul, J. F. & McDonald, M. E. (2005). Development of empirical, geographically specific water quality criteria: a conditional probability analysis Approach. *Journal of the American Water Resources Association*, **41**, 1211–23.

Peck, D. V., Averill, D. K., Herlihy, A. T., *et al.* (2006). Environmental Monitoring and Assessment Program – Surface Waters Western Pilot Study: field operations manual for non-wadeable rivers and streams. EPA 620/R-06/003. Washington, DC: Office of Research and Development, US Environmental Protection Agency.

Peterson, C. G. and Grimm, N. B. (1992). Temporal variation in enrichment effects during periphyton succession in a nitrogen-limited desert stream ecosystem. *Journal of the North American Benthological Society*, **11**, 20–36.

Peterson, C. G. & Stevenson, R. J. (1989). Seasonality in river phytoplankton: multivariate analyses of data from the Ohio River and six Kentucky tributaries. *Hydrobiologia*, **182**, 99–114.

Peterson, C. G. and Stevenson, R. J. (1992). Resistance and recovery of lotic algal communities: importance of disturbance timing, disturbance history, and current. *Ecology*, **73**, 1445–61.

Pielou, E. C. (1984). *The Interpretation of Ecological Data, A Primer on Classification and Ordination*. New York, NY: John Wiley & Sons.

Plafkin, J. L., Barbour, M. T., Porter, K. D., Gross, S. K., and Hughes, R. M. (1989). Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/444/4–89-001, Washington, DC: US EPA Office of Water.

Ponander, K. C., Charles, D. F., and Belton, T. J. (2007). Diatom based TP and TN inference models and indices for monitoring nutrient enrichment of New Jersey streams. *Ecological Indicators*, **7**, 79–93.

Porter, S. D., Mueller, D. K., Spahr, N. E., Munn, M. D., & Dubrovsky, N. M. (2008). Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. *Freshwater Biology*, **53**, 1036–54.

Potapova, M. G. & Charles, D. F. (2002). Benthic diatoms in USA rivers: distributions along spatial and environmental gradients. *Journal of Biogeography*, **29**, 167–87.

Potapova, M. G. & Charles, D. F. (2003). Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition. *Freshwater Biology*, **48**, 1311–28.

Potapova, M. G. & Charles, D. F. (2005). Choice of substrate in algae-based water-quality assessment. *Journal of North American Benthological Society*, **24**, 415–27.

Potapova, M. G. & Charles, D. F. (2007). Diatom metrics for monitoring eutrophication in rivers of the United States. *Ecological Indicators*, **7**, 48–70.

Potapova, M. G., Charles, D. F., Ponader, K. C., and Winter, D. M. (2004). Quantifying species indicator values for trophic diatom indices: a comparison of approaches. *Hydrobiologia*, **517**, 25–41.

Prygiel, J. and Coste, M. (1993). The assessment of water quality in the Artois-Picardie water basin (France) by the use of diatom indices. *Hydrobiologia*, **269/270**, 343–9.

Prygiel, J., Carpentier, P., Almeida, S., *et al.* (2002). Determination of the Biological Diatom Index (IBD NF T 90–354). Results of an intercomparison excercise. *Journal of Applied Phycology*, **14**, 27–39.

Richards, C., Johnson, L. B., and Host, G. E. (1996). Landscape-scale influences on stream habitats and biota. *Canadian Journal of Fisheries and Aquatic Sciences*, **53** (Supplement 1), 295–311.

Rier, S. T. & Stevenson, R. J. (2006). Response of periphytic algae to gradients in nitrogen and phosphorus in streamside mesocosms. *Hydrobiologia*, **561**, 131–47.

Robinson, C. T., Rushforth, S. R., and Liepa, R. A. (1995). Relationship of land use to diatom assemblages of small streams in Latvia. In *A Century of Diatom Research in North America: A Tribute to the Distinguished Careers of Charles W. Reimer and Ruth Patrick*, ed. J. P. Kociolek and M. J. Sullivan, Champaign, IL: Koeltz Scientific Books, pp. 47–59.

Rosati, T. C., Johansen, J. R., and Coburn, M. M. (2003). Cyprinid fishes as samplers of benthic diatom communities in freshwater streams of varying water quality. *Canadian Journal of Fisheries and Aquatic Sciences*, **60**, 117–25.

Rosen, B. H. (1995). Use of periphyton in the development of biocriteria. In *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*, ed. W. S. Davis and T. P. Simon, Boca Raton, FL: Lewis Publishers, pp. 209–15.

Rosen, B. H. & Lowe, R. L. (1984). Physiological and ultrastructural responses of *Cyclotella meneghiniana* (Bacillariophyta) to light intensity and nutrient limitation. *Journal of Phycology*, **20**, 173–93.

Rott, E. (1991). Methodological aspects and perspectives in the use of periphyton for monitoring and protecting rivers. In *Use of Algae for Monitoring Rivers*, ed. B. A. Whitton, E. Rott and G. Friedrich, Innsbruck: Universität Innsbruck, pp. 9–16.

Rott, E. and Pfister, P. (1988). Natural epilithic algal communities in fast-flowing mountain streams and rivers and some man-induced changes. *Verhandlungen Internationale Vereinigung für Theoretische und angewandte Limnologie*, **23**, 1320–4.

Rott, E., Pipp, E., and Pfister, P. (2003). Diatom methods developed for river quality assessment in Austria and a cross-check against numerical trophic indication methods used in Europe. *Algological Studies*, **110**, 91–115.

Round, F. E. (1991). Diatoms in river water-monitoring studies. *Journal of Applied Phycology*, **3**, 129–45.

Rumeau, A., and Coste, M. (1988). Initiation à la systématique des diatomées d'eau douce pour l'utilisation pratique d'un indice diatomique générique. *Bulletin Francais de la Pêche et de la Pisciculture*, **309**, 1–69.

Schaumburg, J., Schranz, C., Hofmann, G., *et al.* (2004). Macrophytes and phytobenthos as indicators of ecological status in German lakes – a contribution to the implementation of the Water Framework Directive. *Limnologica*, **34**, 302–14.

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

Schoeman, F. R. (1976). Diatom indicator groups in the assessment of water quality in the Jukskei-Crocodile River System (Transvaal, Republic of South Africa). *Journal of the Limnological Society of South Africa*, **2**, 21–4.

Schönfelder, I., Gelbrecht, J., Schönfelder, J. & Steinberg, C.E.W. (2002). Relationships between littoral diatoms and their chemical environment in northeastern German lakes and rivers. *Journal of Phycology*, **38**, 66–82.

Shannon, C. F. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, **27**, 37–42.

Sicko-Goad, L., Stoermer, E. F., and Ladewski, B. G. (1977). A morphometric method for correcting phytoplankton cell volume estimates. *Protoplasma*, **93**, 147–63.

Simpson, E. H. (1949). Measurement of diversity. *Nature*, **163**, 688.

Slàdecek, V. (1973). System of water quality from the biological point of view. *Archiv für Hydrobiologie und Ergebnisse Limnologie*, **7**, 1–218.

Smol, J. P., Wolfe, A. P., Birks, H. J. B., *et al.* (2005). Climate-driven regime shifts in the biological communities of arctic lakes. *Proceedings of the National Academy of Sciences of the USA*, **102**, 4397–402.

Soininen, J. (2007). Environmental and spatial control of freshwater diatoms – a review. *Diatom Research*, **22**, 473–90.

Soininen, J. & Niemelä, P. (2002). Inferring the phosphorus levels of rivers from benthic diatoms using weighted averaging. *Archiv für Hydrobiologie*, **154**, 1–18.

Sonneman, J. A., Walsh, C. J., Breen, P. F., and Sharpe, A. K. (2001). Effects of urbanization on streams of the Melbourn region, Victoria, Australia. II. Benthic diatom communities. *Freshwater Biology*, **46**, 553–65.

Squires, L. E., Rushforth, S. R., and Brotherson, J. D. (1979). Algal response to a thermal effluent: study of a power station on the Provo River, Utah, USA. *Hydrobiologia*, **63**, 17–32.

Steinberg, C. and Schiefele, S. (1988). Indication of trophy and pollution in running waters. *Zeitschrift für Wasser-Abwasser Forschung*, **21**, 227–34.

Steinman, A. D., McIntire, C. D., Gregory, S. V., Lamberti, G. V., and Ashkenas, L. (1987). Effect of herbivore type and density on taxonomic structure and physiognomy of algal assemblages in laboratory streams. *Canadian Journal of Fisheries and Aquatic Sciences* **44**, 1640–8.

Stevenson, R. J. (1983). Effects of current and conditions simulating autogenically changing microhabitats on benthic algal immigration. *Ecology*, **64**, 1514–24.

Stevenson, R. J. (1984a). Procedures for mounting algae in a syrup medium. *Transactions of the American Microscopical Society*, **103**, 320–1.

Stevenson, R. J. (1984b). Epilithic and epipelic diatoms in the Sandusky River, with emphasis on species diversity and water quality. *Hydrobiologia*, **114**, 161–75.

Stevenson, R. J. (1990). Benthic algal community dynamics in a stream during and after a spate. *Journal of the North American Benthological Society*, **9**, 277–88.

Stevenson, R. J. (1996). An introduction to algal ecology in freshwater benthic habitats. In *Algal Ecology: Freshwater Benthic Ecosystems*, ed. R. J. Stevenson, M. Bothwell, and R. L. Lowe, San Diego, CA: Academic Press, pp. 3–30.

Stevenson, R. J. (1997). Scale-dependent causal frameworks and the consequences of benthic algal heterogeneity. *Journal of the North American Benthological Society*, **16**, 248–62.

Stevenson, R. J. (1998). Diatom indicators of stream and wetland stressors in a risk management framework. *Environmental Monitoring and Assessment*, **51**, 107–18.

Stevenson, R. J. (2006). Refining diatom indicators for valued ecological attributes and development of water quality criteria. In *Advances in Phycological Studies*, ed. N. Ognjanova-Rumenova and K. Manoylov, Moscow: Pensoft Publishers, pp. 365–83.

Stevenson, R. J. & Bahls, L. L. (1999). Periphyton protocols. In *Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish*, 2nd edn., ed. M. T. Barbour, J. Gerritsen, and B. D. Snyder, EPA 841-B-99–002. United States Environmental Protection Agency, Washington, DC, pp. 6–1 to 6–22.

Stevenson, R. J., Bailey, B. C., Harass, M. C., *et al.* (2004a). Designing data collection for ecological assessments. In *Ecological Assessment of Aquatic Resources: Linking Science to Decision-Making*, ed. M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton, Pensacola, FL: Society of Environmental Toxicology and Chemistry, pp. 55–84.

Stevenson, R. J., Bailey, B. C., Harass, M. C., *et al.* (2004b). Interpreting results of ecological assessments. In *Ecological Assessment of Aquatic Resources: Linking Science to Decision-Making*, ed. M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton, Pensacola, FL: Society of Environmental Toxicology and Chemistry, pp. 85–111.

Stevenson, R. J., Bothwell, M., and Lowe, R. L. (1996). *Algal Ecology: Freshwater Benthic Ecosystems*. San Diego, CA: Academic Press.

Stevenson, R. J., Hill, B. E., Herlihy, A. T., Yuan, L. L., and Norton, S. B. (2008a). Algae–P relationships, thresholds, and frequency distributions guide nutrient criterion development. *Journal of the North American Benthological Society*, **27**, 259–75.

Stevenson, R. J. & Lowe, R. L. (1986). Sampling and interpretation of algal patterns for water quality assessment. In *Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems*, ASTM STP 894, American Society for Testing and Materials Publication, Philadelphia, PA, pp. 118–49.

Stevenson, R. J. & Pan, Y. (1999). Assessing environmental conditions in rivers and streams with diatoms. In *The Diatoms: Applications for the Environmental and Earth Sciences*, ed. E. F. Stoermer and J. P. Smol, Cambridge: Cambridge University Press, pp. 11–40.

Stevenson, R. J., Pan, Y., Manoylov, K., *et al.* (2008b). Development of diatom indicators of ecological conditions for streams of the western United States. *Journal of the North American Benthological Society*, **27**, 1000–16.

Stevenson, R. J. and Peterson, C. G. (1991). Emigration and immigration can be important determinants of benthic diatom assemblages in streams. *Freshwater Biology*, **6**, 295–306.

Stevenson, R. J. & Pinowska, A. (2007). Diatom indicators of ecological conditions in Florida springs. Report to Florida Department of Environmental Protection, Tallahassee, FL.

Stevenson, R. J., Rier, S. T., Riseng, C. M., Schultz, R. E., and Wiley, M. J. (2006). Comparing effects of nutrients on algal biomass in streams in 2 regions with different disturbance regimes and with applications for developing nutrient criteria. *Hydrobiologia*, **561**, 149–65.

Stevenson, R. J. & Smol, J. P. (2002). Use of algae in environmental assessments. In *Freshwater Algae in North America: Classification and Ecology*, ed. J. D. Wehr and R. G. Sheath, San Diego, CA: Academic Press, pp. 775–804

Stevenson, R. J. and Stoermer, E. F. (1981). Quantitative differences between benthic algal communities along a depth gradient in Lake Michigan. *Journal of Phycology*, **17**, 29–36.

Stevenson, R. J. and White, K. D. (1995). A comparison of natural and human determinants of phytoplankton communities in the Kentucky River basin, USA. *Hydrobiologia*, **297**, 201–16.

Stevenson, R. J., Novoveska, L., Riseng, C. M., and Wiley, M. J. (2009). Comparing responses of diatom species composition to natural and anthropogenic factors in streams of glaciated ecoregions. *Nova Hedwigia*, **135**, 1–13.

Stoddard, J. L., Herlihy, A. T., Peck, D. V., *et al.* (2008). A process for creating multimetric indices for large-scale aquatic surveys. *Journal of the North American Benthological Society*, **27**, 878–91.

Stoddard, J. L., Larsen, D. P., Hawkins, C. P., Johnson, R. K., and Norris, R. H. (2006). Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecological Applications*, **16**, 1267–76.

Swanson, C. D. and Bachmann, R. W. (1976). A model of algal exports in some Iowa streams. *Ecology*, **57**, 1076–80.

Taylor, J. C., Prygiel, J,. Vosloo, A., Rey, P. A., d. l., and Rensburg, L., v. (2007). Can diatom-based pollution indices be used for biomonitoring in South Africa? A case study of the Crocodile West and Marico water management area. *Hydrobiologia*, **592**, 455–64.

ter Braak, C. J. F. (1986). Interpreting a hierarchical classification with simple discriminate functions: an ecological example. In *Data Analysis and Informatics*, ed. E. Diday, Amsterdam: North Holland, pp. 11–21.

ter Braak, C. J. F. & Šmilauer, P. (2002). CANOCO Reference manual and CanDraw for Windows User's Guide. Software for Canonical Community Ordination (version 4.5). Biometris, Wageningen and České Budějovice.

ter Braak, C. J. F. & Juggins, S. (1993). Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia*, **269/270**, 485–502.

ter Braak, C. J. F. & van Dam, H. (1989). Inferring pH from diatoms: a comparison of old and new calibration methods. *Hydrobiologia* **178**, 209–23.

Tett, P., Gallegos, C., Kelly, M. G., Hornberger, G. M., and Cosby, B. J. (1978). Relationships among substrate, flow, and benthic microalgal pigment density in the Mechums River, Virginia. *Limnology and Oceanography*, **23**, 785–97.

Tison, J., Park, Y.-S., Coste, M., *et al.* (2007). Predicting diatom reference communities at the French hydrosystem scale: a first step towards the definition of the good ecological status. *Ecological Modelling*, **203**, 99–108.

Tuchman, M. & Stevenson, R. J. (1980). Comparison of clay tile, sterilized rock, and natural substrate diatom communities in a small stream in southeastern Michigan, U.S.A. *Hydrobiologia*, **75**, 73–9.

van Dam, H. (1982). On the use of measures of structure and diversity in applied diatom ecology. *Nova Hedwigia*, **73**, 97–115.

van Dam, H. & Mertens, A. (1993). Diatoms on herbarium macrophytes as indicators for water quality. *Hydrobiologia*, **269–270**: 437–45.

van Dam, H., Mertens, A. & Sinkeldam, J. (1994). A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. *Netherlands Journal of Aquatic Ecology*, **28**, 117–33.

Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., and Cushing, C. E. (1980). The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences*, **37**, 130–7.

Vollenweider, R. A. & Kerekes, J. J. (1981). Background and summary results of the OECD cooperative program on eutrophication. In *Restoration of Inland Lakes and Waters*, Washington, DC: US Environmental Protection Agency, pp. 25–36.

Walley, W. J., Grbović, J., and Džeroski, S. (2001). A reappraisal of saprobic values and indicator weights based on Slovenian river quality data. *Water Research*, **35**, 4285–92.

Wang Y. K., Stevenson R. J., and Metzmeier L. (2005). Development and evaluation of a diatom-based index of Biotic Integrity for the Interior Plateau Ecoregion, USA. *Journal of the North American Benthological Society*, **24** (4), 990–1008.

Watanabe, T., Asai, K., Houki, A., Tanaka, S., and Hizuka, T. (1986). Saprophilous and eurysaprobic diatom taxa to organic water pollution and diatom assemblage index (DAIpo). *Diatom*, **2**, 23–73.

Weber, C. I. (1973). Recent developments in the measurement of the response of plankton and periphyton to changes in their environment. In *Bioassay Techniques and Environmental Chemistry*, ed. G. Glass, Ann Arbor, MI: Ann Arbor Science Publishers, pp. 119–38.

Weilhoefer, C. L. & Pan, Y. (2007). A comparison of periphyton assemblages generated by two sampling protocols. *Journal of North American Benthological Society*, **26**, 308–18.

Whitton, B. A. & Kelly, M. G. (1995). Use of algae and other plants for monitoring rivers. *Australian Journal of Ecology*, **20**, 45–56.

Whitton, B. A., Rott, E., and Friedrich, G. (ed.) (1991). *Use of Algae for Monitoring Rivers*, Innsbruck: Universitat Innsbruck.

Winter, J. G. & Duthie, H. C. (2000). Epilithic diatoms as indicators of stream total N and total P concentration. *Journal of the North American Benthological Society*, **19**, 32–49.

Wolda, H. (1981). Similarity indices, sample size and diversity. *Oecologia*, **50**, 296–302.

Wu, J. T. (1999). A generic index of diatom assemblages as bioindicator of pollution in the Keelung River of Taiwan. *Hydrobiologia*, **397**, 79–87.

Wunsam, S., Cattaneo, A., and Bourassa, N. (2002). Comparing diatom species, genera and size in biomonitoring: a case study from streams in the Laurentians (Québec, Canada). *Freshwater Biology*, **47**, 325–40.

Yallop, M., Hirst, H., Kelly, M., *et al.* (2006). Validation of ecological status concepts in UK rivers using historic diatom samples. *Aquatic Botany* **90**, 289–95.

Zelinka, M. & Marvan, P. (1961). Zur Präzisierung der biologischen Klassifikation des Reinheit fliessender Gewässer. *Archiv für Hydrobiologie*, **57**, 389–407.