River macrophyte sampling: methodologies and variability

Report: SC070051/R1

Integrated catchment science programme

Evidence Directorate

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned and funded by the Environment Agency.

Published by:
Environment Agency, Rio House, Waterside Drive,
Aztec West, Almondsbury, Bristol, BS32 4UD
Tel: 01454 624400 Fax: 01454 624409
www.environment-agency.gov.uk

ISBN: 978-1-84911-139-3

Author(s):
Allan Pentecost (Freshwater Biological Association)
Nigel Willby (University of Stirling)
Jo-Anne Pitt (Environment Agency)

Dissemination Status:
Released to all regions
Publicly available

Keywords:
macrophyte, metric, variability

Research Contractor:
The Freshwater Biological Association
The Ferry Landing
Far Sawrey, Ambleside
Cumbria, LA22 0LP, UK
Tel: +44 (0)1539 442468

Environment Agency's Project Manager:
Veronique Adriaenssens, Evidence Directorate

Collaborator(s):
Countryside Council for Wales
Natural England
Scottish Environment Protection Agency

Project Number:
SC070051/R1

Product Code:
SCHO1109BRHI-E-P

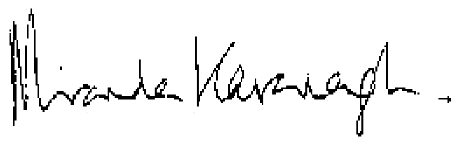
Evidence at the Environment Agency

Evidence underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us, helps us to develop tools and techniques to monitor and manage our environment as efficiently and effectively as possible. It also helps us to understand how the environment is change and to identify what the future pressures may be.

The work of the Environment Agency's Evidence Directorate is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The Research & Innovation programme focuses on four main areas of activity:

- **Setting the agenda**, by informing our evidence-based policies, advisory and regulatory roles;
- **Maintaining scientific credibility**, by ensuring that our programmes and projects are fit for purpose and executed according to international standards;
- **Carrying out research**, either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.



Miranda Kavanagh

Director of Evidence

Executive summary

Current sampling strategies and variability studies for aquatic river macrophytes across the EU are reviewed in this report. Macrophytes are water plants that are visible to the naked eye. The spatio-temporal variation of aquatic macrophytes in rivers is notoriously high, and this combined with the difficult sampling environment is likely to lead to greater sampling errors than in surveys of terrestrial vegetation. There are at present too few detailed studies to quantify the sampling errors associated with surveys of aquatic vegetation. Particular difficulties identified are variable river depth, uncertainty regarding the aquatic/terrestrial boundary, difficulties in the identification of taxa within critical plant groups, such as bryophytes and the genus *Ranunculus*, and variable but potentially low detection rates.

The optimum size of the sampling unit is also reviewed here. Surveyors use a reach length ranging from two to 500-m, with 100-m being the most popular. These lengths are not justified in terms of species/area relationships as they tend to be in terrestrial ecology, and there is a lack of a theoretical basis for the reach lengths used. Reach length could be varied to standardise the area surveyed, but this may not be appropriate given the dependency of river plants on linear habitat features. Reach lengths fixed at 100 m have advantages (for example in facilitating international comparisons of results, and in lowland areas where land use mosaics are on a scale of one km or less), but may limit the comparability of surveys when habitat features vary over large scales (greater than 100-m), as in large lowland rivers. In upland rivers that vary physically over tens of metres, a fixed 100-m survey length may be less problematic and uncommon habitat features, such as debris dams or waterfalls, are likely to account for a small proportion of the total survey length. In gemorphological terms, it is arguable that the length of river assessed should be a fixed multiple of channel width to reflect the changing periodicity of major features, such as riffle and pool sequences. In lowland rivers, a longer survey length would also be consistent with the increasing size of individual plants and the beds that they form, compared to the vegetation of small upland streams. The significance of the effect of survey length on survey results is very dependent on the purposes of a survey and how the data is to be interpreted.

The number of surveys per water body varies with the purpose of the survey but is low in most EU countries. Relatively little research has been undertaken on levels of species turnover between surveys in rivers of different types, or on the numbers of surveys that would be required to best characterise the vegetation.

Variability of a range of factors is reviewed. There are little data on the variability of the primary metrics such as species cover and richness, but some simulation exercises indicate that, as far as the important water quality metrics such as MTR (Mean Trophic Rank) are concerned, misidentification will in general lead to larger errors in the final river classification than inaccurate cover estimates. There are a few studies on the variability in derived metrics such as MTR. They indicate that the observer precision measured as the coefficient of variation is about five per cent. Temporal variability over season and year indicate more variability than observer error, but there are large differences between the few studies available so no general conclusions can be made. Further studies of variability are recommended, with sufficient replication to account for spatio-temporal interaction, observer error, and differences in the physical nature of rivers from source to sea.

Acknowledgements

I am grateful to the following individuals for the provision of current information on this topic: Veronique Adriaenssens (Environment Agency, U.K.), Annette Baatrup-Pedersen (NERI, Denmark), Sebastian Birk (University of Duisburg-Essen, Germany), Bill Brierley (Environment Agency, U.K.), Joe Caffrey (U.K.), Mike Dobson (Freshwater Biological Association, U.K.), Willie Duncan (SEPA), Frauke Ecke (Swedish Intercalibration Group), Laura Grinberga (Latvian Intercalibration Group), S. Hellsten (Finnish Intercalibration Group), George Hinton (Natural England, U.K.), Ian Johnson (DOE, Northern Ireland), Chris Mainstone (Natural England, U.K.), Simon Pawley (FBA, U.K.), Susi Schneider (Norwegian Institute of Water Research), Zofija Sinkeviciene (Botanical Institute, Lithuania), Roger Sweeting (Freshwater Biological Association, U.K.),and Nigel Willby (University of Stirling, U.K.).

Contents

List of figures

List of tables

1 Introduction

The aim of this report is to provide current information on methods used in river macrophyte surveys within the European Union (EU) with respect to their variability. Macrophytes[1] are an important component of European river systems, providing a source of food and shelter for a wide range of organisms and adding aesthetically to wetland landscapes. In some situations their influence may be undesirable, such as the impedance of water flow leading to flooding. Their decay *en masse* may also contribute to river deoxygenation and fish death in nutrient-enriched rivers. Aquatic macrophytes are known to respond to a large number of environmental factors. The more significant of these are the rate and variability of water flow, alkalinity, substrate, shading, and nutrient concentrations but many other factors are likely to have indirect effects (Barendregt & Bio, 2003; Lacoul & Freedman, 2006). The responsiveness of macrophytes to factors that are strongly influenced by anthropogenic activity can make them useful as indicators of river quality, although co-variation with other factors may render identification of specific causes, such as phosphorus-enrichment, difficult.

It is known, for example, that *Scapania undulata* grows preferentially in water that is low in dissolved nutrients and high in oxygen while others, such as the moss *Leptodictyon riparium,* grows well in nutrient-enriched sites such as sewage filter beds. Thus, it has been possible to rank aquatic macrophytes along gradients of different chemical and physical conditions so that a particular species is assigned a 'score' related to the conditions under which it is most commonly found. Such scores can be assembled to provide a total score for a particular site that should integrate information about the environment as experienced by the vegetation. The scores generated in this way are often referred to as metrics or indices (although other metrics, such as total cover or species richness, may be derived from the same primary biological data). Metrics derived in this way have been widely used to assess changes upstream and downstream of point sources, for example for the Urban Waste Water Treatment Directive (UWWTD). Subsequent to the Water Framework Directive (WFD), benchmarking of observed metrics against the values to be expected at unimpacted reference sites has become the norm in Europe.

A full macrophyte survey can be broken down into a sequence of four operations: 1) a pilot survey to decide how and where the study is to be conducted; 2) the recording of macrophytes in the chosen area/length of river; 3) assembly of the data; and 4) manipulation of those data to generate the desired metric(s).

Since river systems are usually complex, extensive and sometimes difficult to access, surveys will normally be undertaken over short lengths. Entire systems are rarely investigated. This leads to obvious problems in choosing 'representative reaches'. River ecologists recognise this problem and usually sacrifice statistical rigour, which requires

[1] Macrophytes are larger plants of freshwater which are easily seen with the naked eye, including all aquatic vascular plants, bryophytes, stoneworts (Characeae) and macro-algal growths.

some form of random sampling, and substitute a practical sampling scheme based on ‘common sense’. A number of schemes have been suggested to provide representative samples such as a fixed proportion of the river length at regular intervals. Some form of stratified sampling based upon river hierarchies has also been suggested such as sampling within a specified river basin, with a fixed number of sites on rivers of particular order (Haury & Muller, 1991). This method offers many advantages but often requires a large number of sites.

The most common form of sampling involves a fixed length of river whose location is chosen to provide an adequate summary of a larger reach or occasionally of the entire river. Certain caveats are usually introduced, such as avoidance of bridges, dams, weirs or other modifying structures, or of waterfalls or other natural hazards. In some cases, an attempt is made to ‘pair’ two reaches above and below a presumed point source of pollution, such as a sewage treatment works. The actual length chosen, however, is not based upon particular ecological principles, as will be shown later.

The recording of aquatic macrophytes is less easily achieved than the recording of terrestrial vegetation. While this is well recognised, much less is known about the reliability of aquatic macrophyte recording. Most ecologists use an estimate of percentage cover to reflect the amount of a particular plant species present (Kershaw & Looney, 1985). This is potentially superior to frequency or presence/absence as it conveys more information, and can be used to generate a wider range of metrics. The main disadvantage is the greater amount of time and therefore cost involved in completing such a survey. Estimates of cover are made by eye. While this is undesirable in terms of both accuracy and precision, there is no alternative using current technology if a large number of sites are to be investigated by small teams of recorders. Equal if not greater attention needs to be paid to misidentification of species, and the overlooking of species. Some widespread macrophyte genera such as *Callitriche* and *Ranunculus* contain critical species requiring substantial expertise to identify. Inexperienced recorders may misidentify taxa more frequently which can potentially result in an error in a metric that is far greater than one resulting from even a very large error in cover estimation. The true effect of such errors will, however, depend on the differences in ranks between the confusable species. If several closely related members of the same genus have similar ranks the effect of misidentification will be much less than that of misidentifying or failing to detect species with contrasting ranks. The last two stages of a survey, the assembly of data and use of metrics for biological assessment, are not covered in this report.

Sampling methods do not differ greatly across the EU. All involve wading into the water where conditions allow and following a standard, usually zig-zag pattern upstream along the length of the reach and are detailed in the water quality guidance standard for aquatic macrophytes (BSI, 2003). Various designs of water-viewer are used to reduce light reflection and are made ‘in house’. These are used routinely in the UK, Germany and other parts of the EU (Baatrup-Pedersen *et al*., 2006). Boats, or occasionally diving, are used where the water is too deep to wade, and macrophyte samplers such as grapnels (UK) or rakes (Germany) are used where necessary. Underwater cameras have been used on occasion but not routinely.

2 Sampling methods for aquatic macrophytes

2.1 Methods used in published studies

The quantitative study of aquatic macrophytes in Europe is a relatively recent development dating to the late 1960s. Prior to this, surveys were usually of the 'presence-absence' type exemplified by botanical recording and were rarely confined to defined lengths or areas of river. The detailed mapping and quantitative surveys carried out in Britain by Butcher in the 1920s (Butcher, 1933) are an exception.

A wide range of recording and sampling methods have been used, as dictated by the nature of the river and the particular requirements of the study. A list of some of these, country by country, is shown in Table 2.1. Survey lengths range from two to 500 metres but the length employed is rarely justified in the text. In 13 out of 15 studies, taxa level cover values form the raw data from which is derived a range metrics such as IBMR (Indice Biologique Macrophytique Rivière; Haury *et al*., 2002) and MTR (Mean Trophic Rank; Holmes *et al*., 1999). In only one case (Denmark; Baatrup-Pedersen *et al*. 2002) was a fixed area examined. In the remainder, the actual area of study is not given. This could lead to problems in comparing sites (in terms of species richness, for example) where the channel width is very different. However, the intention is normally to compare rivers of a similar type (for example, a lowland river versus a river of the same type in reference condition), rather than a lowland river versus an upland one. It also seems likely in a river, that the longitudinal and lateral axes of a reach cannot be considered equivalent and therefore the linear length surveyed is probably more critical than the area.

Table 2.1 Examples of reach lengths and derived metrics used in published river macrophyte studies across Europe.

Country	Reach length m.	Metrics	Notes	Reference
DENMARK	200	Cover% for RHS	25 x25 cm plots examined in transects every 10 m. Study of weed cutting in rivers	Baatrup-Pedersen *et al.* (2002)
FINLAND	5 m strips along one bank	Cover%	Used alternating strips along one bank. One of earliest quantitative studies	Sirjola (1969)
ESTONIA	100	Cover%, IBMR, MTR	Hierarchical study of a river basin	Springe & Sandin (2004)
FRANCE	50	Cover%, presence	Stratified sample of upper and lower river reach	Haury & Muller (1991)
FRANCE	2	Cover%	At 25-50 m intervals along floodplain rivers	Bornette (2001)
GERMANY	50	Cover%	Long term changes in floodplain river	Wiegleb *et al*. (1989)
GERMANY	100	I-IV scale of abundance	Looking at macrophyte structure in relation to flow	Passauer *et al*. (2002)
POLAND	100	Cover%, IBMR, MTR	Comprehensive study, part of STAR project	Szoszkiewicz *et al*. (2004); Staniszewski *et al*. (2006)
UK	500	Presence/absence	Early use of fixed length. Length chosen not justified	Whitton & Buckmaster (1970)
UK	500	Cover%	Detailed study of N. England river	Holmes & Whitton (1977)
UK	500	Presence/absence	Spaced 4 km apart Conservation evaluation	Wilkinson *et al*. (1998)
UK	500	Cover %	Natural England, CCW, SNH	JNCC (2005)
PAN EUROPE	500	Cover%, MTR	Diversity of upland and lowland rivers	O'Hare *et al*. (2006)
PAN EUROPE	100	IBMR, MTR, GRI	Intercalibration study	Birk *et al*. (2006)
PAN EUROPE	100	Cover%, IBMR, MTR	Study of unimpacted sites	Baatrup-Pedersen *et al*. (2006)
PAN EUROPE	100	Cover%, MTR variant	Nutrient effects	Johnson *et al*. (2007)

2.2 Aquatic macrophyte monitoring schemes in EU member states

The Water Framework Directive (WFD) Annex V (Directive 2000/60/EC) requires the monitoring of the composition and abundance of the aquatic flora of rivers and assignment of ecological status to one of five classes. The classification is based upon an Ecological Quality Ratio (EQR) which is the ratio of a derived metric (or combination of metrics) from the sample site to metric(s) obtained from an appropriate unimpacted ‘reference site’. The WFD requires estimates of the level of confidence and precision in the results but it does not give *a priori* levels of expected precision or confidence. The UK Technical Advisory Group (UKTAG) have indicated the desirability of classification schemes for biological quality elements that deliver 95 per cent confidence of class when samples are located in the middle of that class (UKTAG, 2007).

Information on current monitoring methods used in the UK and surrounding ecoregions, as defined in the WFD, is shown in Table 2.2. Ecoregions covering south-eastern Europe and the Mediterranean have not been included owing to their substantially different climate and vegetation. Of the sixteen states included, eleven use a reach length of 100-m either exclusively or partially, and one uses either 50-m, 200-m or 500-m. Some states have yet to indicate the length of reach they will use.

Table 2.2 Reach lengths and derived metrics used in different member states for the purposes of routine monitoring.

Country	Reach length m	Metrics	Notes	Reference
AUSTRIA	100	Ecological Integrity Assessment [no details, being developed]	500 m used for large rivers. Emergent macrophytes also sampled. Uses a total of four metrics	Birk *et al*. (2007)
BELGIUM	100	Integrates four metrics	Plans to use IBMR, no variability studies	Birk *et al*. (2007)
DENMARK	100	Related to MTR	Variability not yet studied	A. Baatrup-Pedersen (pers. comm.)
EIRE		No information		
FINLAND	200	Not developed	Plan to use the Swedish Nielson 7-point scale scheme	S. Hellsten (pers. comm.)
FRANCE	100 with a minimum sampling area of 50 m^2	IBMR	Similar to MTR but with a coarser scale	Haury *et al*. (2002)

GERMANY	100 (max)	A reference index system similar to MTR	Variability studies are in progress	Schaumberg *et al*. (2004)
LATVIA	100	STAR protocol	No variability study reported	L. Grinberga (pers. comm.)
LITHUANIA	100	Not known	Used in rivers with catchments up to 10,000 km^2, otherwise longer	Z. Sinkeviciene (pers. comm.)
NETHERLANDS	50	AMOEBE	Under development. Emergents also recorded, but not bryophytes. Usually combine 3x50 m reaches	Birk *et al*. (2007)
NORWAY	50-100	Currently no official standard	In development stage	S. Schneider (pers. comm.)
POLAND	100	MIR (Macrophyte Index for Rivers). Related to MTR	500 m length used for large rivers	Birk *et al*. (2007)
RUSSIA	100-300 (2000)	No information	Included for comparison. Non EU member.	L. Grinberga (pers comm.)
SWEDEN	-	No current monitoring	In development stage	
UK (England, Wales & N.Ireland)	100	MTR	50 or 500 m used on some reaches. Variability assessed but mostly on a subjective basis	Holmes *et al*. (1999)
UK (SCOTLAND)	100	LEAFPACS	3 to 5 x100 m reaches surveyed. 'Models' expected metrics on site-specific basis using abiotic variables. Variability studies in progress (2008)	Willby *et al*. (2009)

2.3 Reach length and its justification in ecological research and regional surveying

Few macrophyte surveys explain why a particular length of reach has been chosen. The most frequently used reach length of 100-m appears to originate in the MTR described by Holmes (1991) but there is no explanation for why this length was chosen. The national methods for river macrophyte assessment of many other EU member states use the same sample length. In some cases these countries have largely adopted the MTR method with small revisions (such as Poland), while in other cases 100-m is simply deemed a practical size of reach to survey. Not all monitoring bodies in the UK use this reach length. The national conservation agencies use a length of 500-m, as recommended by the Common Standards Monitoring Guidance for Rivers (JNCC, 2005). A 500 m reach length has also been suggested for sampling larger rivers, for example in Austria and Poland. However, in the UK, the MTR uses 50-m if a river is greater than 50-m wide. A simple conclusion from the range of approaches available is that one size cannot fit all purposes.

Contrasting lengths could be justified on the basis of area, since large rivers give a greater sampling area than small rivers over the same length. However, the general approach on large continental rivers is to use longer survey lengths. Some ecologists such as Frisell *et al*. (1986) argue for a variable reach length for stream habitat classifications. They suggest a range from tens of metres in steep streams to hundreds of metres in fifth order or larger rivers. The basis for this lies in the periodicity of physical habitat features generated by basic fluvio-geomorphological processes; in very small rivers, riffle-pool sequences may occur over tens of metres while in large lowland rivers the same features may occur over scales of hundreds of metres. As a rough rule of thumb, riffle-pool sequences will occur at a frequency equivalent to five to seven channel widths, with the sequence length increasing with respect to channel width as slope and particle size decreases (Richards, 1976). To ensure a roughly equivalent level of coarse-scale physical habitat variability, the length of reach surveyed should increase with stream order. Thus, a standard sample length will grossly under-represent physical habitat variability on large rivers compared to small rivers. Some further justification can be found in the body sizes the taxa encountered. Steep, low order streams tend to be dominated by bryophytes whose individual biomass, and individual cover, may be several orders of magnitude lower than that of rooted vascular macrophytes in large lowland rivers. The BSI (2003) does not stipulate a fixed length but recommends a study of species richness versus reach length. It states that the length is to *'reflect adequately the diversity of plant species characteristic for the [particular] ecological type of river*'.

Another consideration in the choice of survey length is the proportion of the total length of a river or water body that will yield a representative sample. For example, reaches of 500-m at a distance of five-km apart will sample 10 per cent of the river system and some publications advocate such a method (Wilkinson *et al*., 1998; JNCC, 2005; EA Operational Guidance Manual, 2007). However, there appears to be no information on the proportion of a river that requires sampling – presumably because rivers are physically heterogeneous, especially in their upper reaches, and a systematic treatment as described above will rarely be adequate throughout a river's length. Wiegleb (1989) argued using his knowledge of European rivers, and a semivariogram approach to vegetation homogeneity (Palmer, 1988), that a reach of about 50-m would be appropriate for macrophyte sampling. (A semivariogram is a graph that relates the variance of a factor to scale. In the case of rivers, the scale may range from a few cm to greater than100-km, and the factor may be the river gradient, or vegetation cover, for example. Analysis of semivariograms allows factor variability to be determined over wide ranges in scale. Unfortunately, the above estimate of Wiegleb (1989) is too vague to discount 100-m or even 500-m reaches as it is not based on a full quantitative analysis. Survey length is directly related to spatial frequency of surveys (the number of surveys per water body on a given date) since survey length will generally decrease as survey frequency increases. If the species pool is large or the turnover between sites is high, more surveys, rather than longer length surveys, may be needed for some river types if the aim is to characterise the vegetation.

Species richness almost always correlates positively to sample area. In terrestrial ecology species-area curves are often used to determine the best sample size based on nested quadrats. This appears not to have been done with aquatic vegetation. Data reported by Holmes (1980) can be used as the basis for a simple analysis, although this does not reveal the number of reaches or optimum length of reach needed to sample a particular river, river type or rivers in general. Holmes (1980) reports the results of surveys of 234 reaches of one km, within which was nested a 500-m reach, and within that a 100-m reach. The numbers of taxa relate to the channels and bank and cover those on the JNCC recording list. On average a one-km reach contained 20 per cent more taxa than a 500-m reach and 68 per cent more taxa than a 100-m reach. The 500-m reach, on average, contained 43 per cent more taxa than the 100-m reach (Figure 2.1).

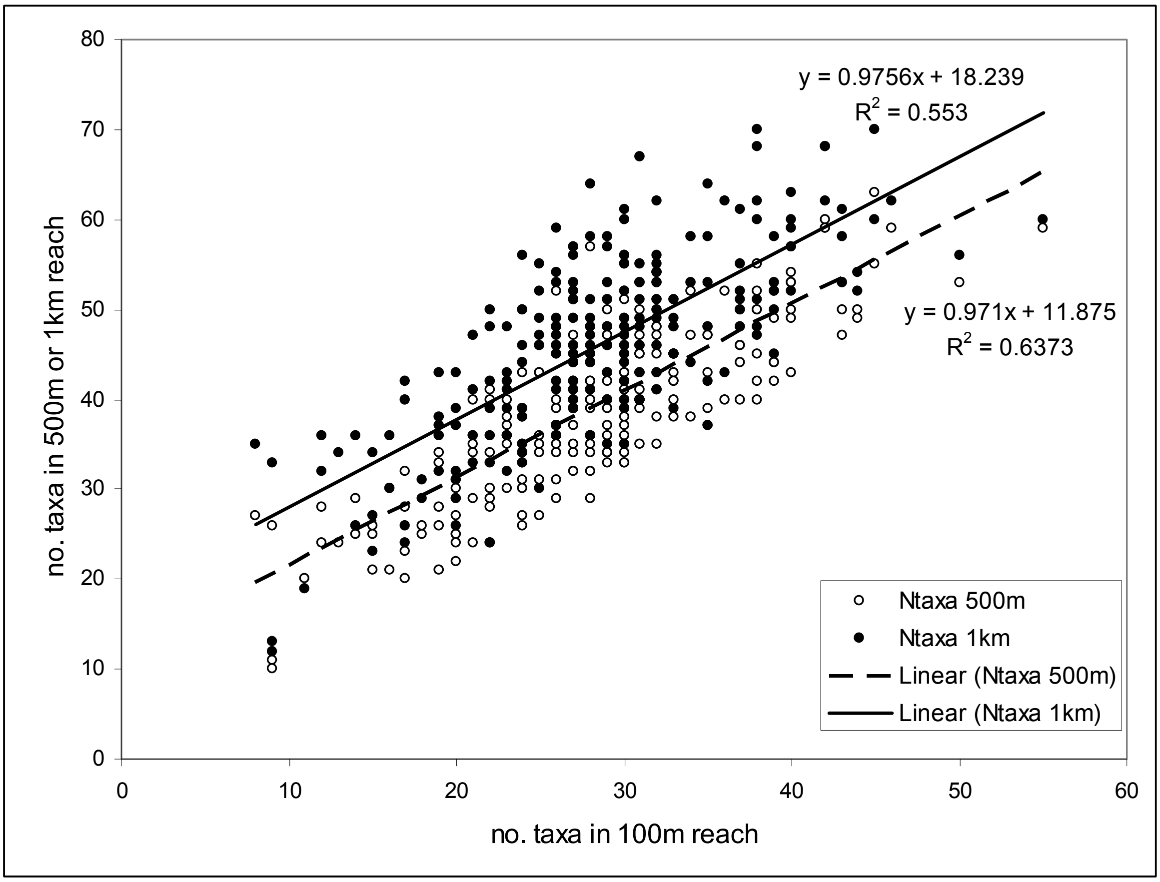


Figure 2.1 Relationship between numbers of taxa in different lengths of recording reach nested within a single site. Data from Holmes (1980).

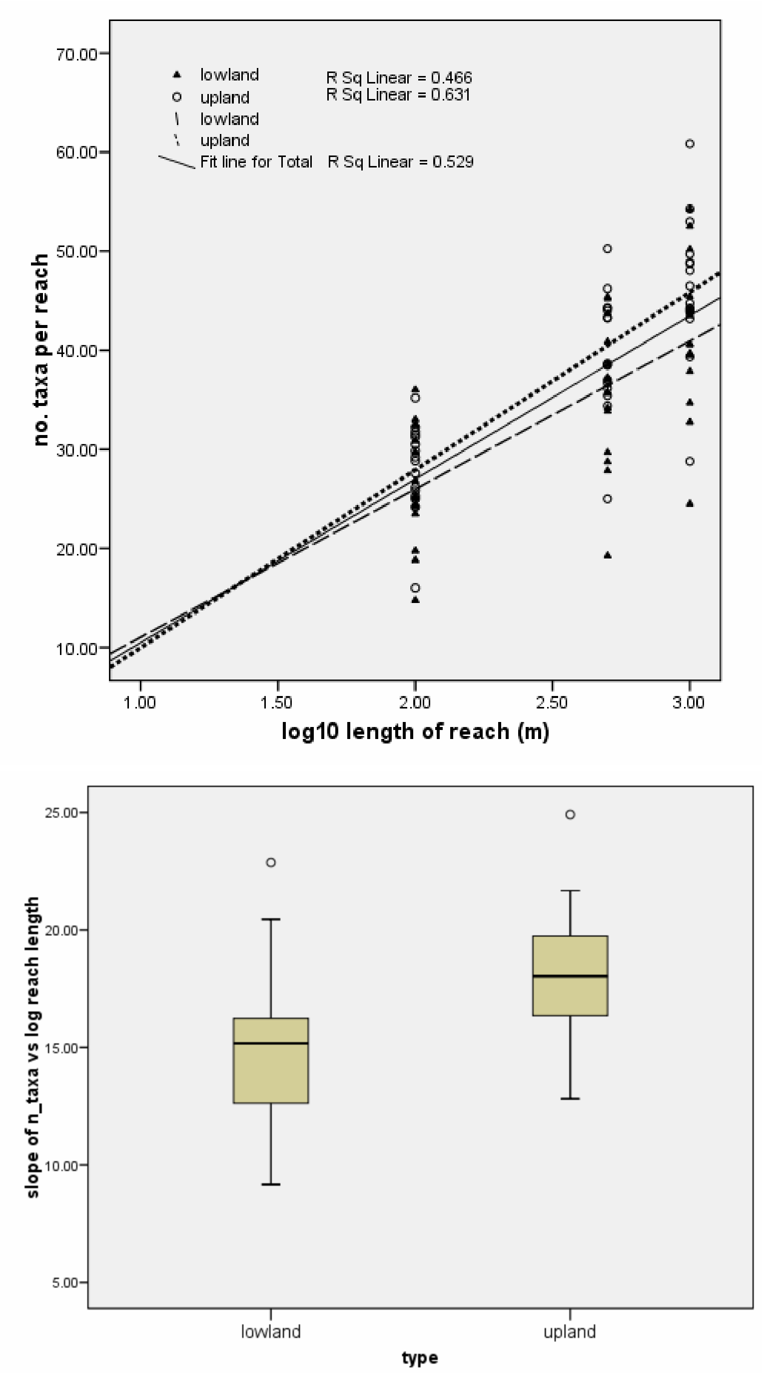


Figure 2.2 Taxa versus log reach length relationships, showing differentiation in slope between upland and lowland rivers.

A more detailed assessment of the data reveals few trends at the level of individual rivers, although the ratio of numbers of species in 1km:100m shows a declining trend from source to mouth in a small number of well-sampled lowland rivers. Species-area relationships normally follow a log linear form. In terms of the relationship between numbers of taxa and sample length there is a strongly significant relationship with a typical slope of 16 (Figure 2.2). Subdivision of rivers into upland and lowland categories (based on a threshold of 200-m altitude at source) reveals that the slope of the richness versus length is significantly steeper ($p = 0.02$) in upland than lowland rivers (18 versus 15 respectively, Figure 2.2). This indicates that, on average, a 100-m reach of lowland river will sample a slightly higher proportion of the species found in one km of river, than in an upland river.

Based upon a large series of 100-m and 500-m MTR values in the same stretch of river, Dawson *et al*. (1999) have shown statistically that there is little difference between them. This suggests that at least for the rivers surveyed, data from a 100-m reach length is adequate for generating unbiased values of the MTR metric. However, if the aim is to identify reaches supporting scarce or uncommon taxa with a low probability of occurring in short lengths, reaches of 500-m or longer are better. This emphasises the underlying importance of recognising the purpose of a survey and adopting a survey strategy fit for purpose.

One argument in favour of shorter survey reaches relates to homogeneity. In catchments of high drainage density, reaches exceeding one km might include inputs from lower order streams with the potential for large nutrient loadings or other pollutants part way down the reach. This is undesirable and could influence macrophyte growth profoundly. In lowland reaches, artificial structures are common and are easier to avoid in reaches of short length, unless they are so frequent and widely distributed that to avoid them would make the survey unrepresentative of the water body. Natural hazards such as waterfalls and log jams and artificial structures such as weirs also lead to difficulty if they fall within surveys reaches of longer length. The mosaic-like nature of lowland Britain is such that adjacent land use often changes dramatically over distances of the order of one km. Forestry and fields holding pasture or livestock often influence river systems through shading/exposure, cattle fouling, diffuse nutrient inputs and trampling. It is clearly much easier to site a 100-m stretch of river in a reasonably homogeneous adjacent environment (whether from the perspective of key environmental factors, such as underlying geology, or in terms of the distribution of pressures) than it is to position a 500-m stretch. However, most of the need for homogeneity of reaches has arisen from pair-wise comparisons of reaches above and below point sources where it is desirable to have reaches that are comparable in other respects. For assessment of ecological status homogeneity is less critical, indeed the loss of reach scale physical heterogeneity may be a function of man-made pressures. If high physical heterogeneity is the natural state for reaches within a water body this will be reflected in the values of biological metrics to which a test reach is to be compared.

Choice of reach length is likely to rest, at least partially, on practicality issues, such as time and cost. In some cases, for example, it may simply be considered more cost-effective to collect data from a smaller number of longer reaches, a decision which has no ecological basis.

2.4 Frequency and timing of sampling

How is the frequency of sampling sites determined within the year and between years? Table 2.3 provides information on the timing of sampling for some EU member states. All sampling periods cover the summer growth period with the longest periods (five months) used by Austria and the UK (Natural England and Countryside Council for Wales). Longer periods can be justified in countries subjected to extreme variation in climatic conditions imposed by topography or wide latitudinal range. The shortest sampling periods are used by Germany, Latvia and Lithuania (three months), while some states, such as France, have not specified a sampling period. Other recommendations with respect to season are two samplings in the year, one early and the other late to account for early- or late-growing macrophytes (BSI, 2003). This can have advantages as flowering may be required to confirm plant identification. Late sampling may also have disadvantages, such as algal smothering, making cover estimation and identification of smothered plants difficult. Although sampling twice a year should give a better assessment of the vegetation, once started it would need to be continued over succeeding years if trends are to be monitored. In the UK for example, there is a three-year rolling sample scheme (each site is visited once every three years) between June and September.

Sampling frequency falls in the range of once per year to once per six years and averages about once per two or three years (Table 2.3). The WFD stipulates sampling once every three years and both Germany and Austria currently adhere to this timing. Frequency has obvious cost implications; considering the high natural variability that can occur between macrophyte communities in successive years, sampling more regularly than three years may be desirable if the focus is on more variable traits such as cover. Generally, metrics related to community structure will be more stable over time since the vegetation is likely to bear the signature of short-term predictable fluctuations (such as in flow) by favouring species with superior performance under naturally variable conditions. Factors that cause significant unpredictable changes in macrophyte composition (such as extreme floods, major anthropogenic disturbances) will occur at a much lower frequency than annually.

Table 2.3 Timing and number of macrophyte survey sites in water bodies in the EU.

Country or region	Sampling period recommended	Between-year sampling strategies	Sites per water body recommended	Reference
AUSTRIA	May-September, lowlands before uplands, with June-August optimal	Once in three years [WFD]	One or more	Birk (2007)
DENMARK	June-September	Once per year	Random site selection, no specified number	Baatrup-Pedersen (pers. comm.)
FRANCE	Differs according to latitude	Once or twice per year	Variable and under review	Birk (2007)
GERMANY	July-September	Once every three years [WFD]	One, but under review	S. Birk (pers. comm..)
LATVIA	July-September	Every three to five years	Once per year is normal	L. Grinberga (pers comm.)
LITHUANIA	July-September	Once per year	Several sites per river	Z. Sinkeviciene (pers comm.)
NETHERLANDS	No data	-	-	Birk (2007)
NORWAY	No details, later at higher latitudes	-	-	Birk (2007)
UK ENVIRONMENT AGENCY [MTR]	Mid June-mid September	Four surveys over three years, and at least one survey every three years	One to three dependent upon size of water body	Holmes *et al*. (1999); V. Adriaenssens (pers. comm.)
UK NATURAL ENGLAND & CCW	May-late September	Sampling every six years	Usually targeted to specific sites such as SSSI's	BSI (2003); C. Mainstone (pers. comm.)
U.K. SEPA	June-September	-	Three sites per water body planned	Willby *et al*. (2009)

2.5 Number of sites per water body

The number of sites per water body has been given some attention but there are no hard and fast rules at present. The WFD defines a 'body of surface water' in Article 2 as '*a discrete and significant element of surface water such as a lake, a reservoir, a stream, river or canal, part of a stream, river or canal [or] a transitional water*'. Clearly the size of a water body should influence the number of sites required, although this will also depend on its characteristics. For example, a reasonably homogeneous stretch of water of roughly equal gradient and width should require fewer sites than one that is heterogeneous with wide-ranging gradients. The LEAFPACS method suggests three to five sites per water body for ecological quality assessment for rivers based on macrophytes. The current situation, as indicated in Table 2.3, shows a wide range of practices from one to six sample sites per water body.

In the UK, river water bodies are defined by the WFD system A typology which generates around 10,000 individual water bodies using a typology based on catchment geology, average catchment elevation and catchment area. Water body length increases downstream, with small upland rivers tending to have water bodies one to five km in length compared to 20-30 km in large lowland rivers.

The question of how many samples to take per water body depends on the purposes of data collection. Targeted sampling approaches, such as the MTR, seek to maximise between-site variability by comparing samples taken above and below point sources. The WFD seeks to classify a water body based on the collection of a representative number of samples that reflect the variability inherent within that water body, whether due to natural factors, or localised impacts. Sites are classified on a site-specific basis in relation to the physical attributes (slope, distance from source, altitude and source altitude) relevant to that site. Consequently, there may be no good reason why the number of samples in a physically heterogeneous but short upland river water body should differ from a long but more homogenous large lowland river. However, if the aim is to characterise the vegetation, for example to assess its condition or describe the composition, some initial appreciation of the potential species pool and the degree of turnover between sites may be necessary to justify the number of reaches to survey.

3 Accounting for variability and quality control

3.1 Estimation of primary metrics

3.1.1 Estimation of cover

All current methods for the assessment of aquatic macrophytes, such as LEAFPACS, CSM and MTR, rely on taxa level cover values as the primary biological data. The other primary metric that finds some use is species or taxa richness. Cover is considered by plant ecologists as the most convenient unit for vegetation classification (Kershaw & Looney, 1985). It has theoretical advantages over presence/absence and frequency but it is not the best metric available. Plant biomass is of more value but is too difficult to sample and estimate. It is also destructive. In aquatic vegetation of standing waters the degree of vertical development through the water column may be as important as spatial extent. This is often reflected in the use of PVI (Plant Volume Infested or Inhabited) as a measure of abundance in preference to simple cover estimates. Due to the constraints of flow on the vertical development of vegetation in running waters it is normal to view river vegetation in two-dimensional terms. Cover estimation in river surveys is done by eye. Again this is far from ideal but alternatives using current technology are extremely time-consuming. The limitations of eye-estimation are well documented in terrestrial vegetation. For example, Greig-Smith (1957) found that terrestrial cover estimates differed from the group mean by up to 25 per cent among eight individuals. The author found his own estimate of cover was in error by an average of 16 per cent from a known area of vegetation.

Observer errors are divided into intra-observer (individual variability/errors) and inter-observer (differences between two or more observers). Intra-observer variability is a well-studied phenomenon and observations carried out by individuals for a long period, say three to four hours, are known to be subject to ‘observer drift’. This is a change in recording quality over time due to tiredness and unconscious reclassification of the vegetation patterns. Inter-observer variability is due to differences in individual perceptions of the study, experience and attitudes towards the work at hand. Its effects can be accommodated (for example by revising or merging cover classes) providing all observers record the same reach and are able to discuss the reasons for the resulting differences in measurement (Ruxton & Colegrave, 2003). However, inter-observer variability is a reality that cannot be removed.

With respect to river macrophytes, little appears to have been published on cover variability although original data must exist. Staniszewski *et al*. (2006) report on the variability of MTR values between sets of three observers recording the same reach but do not tabulate the original cover values. Holmes *et al*. (1999) report that a difference in cover of greater than15 per cent between two independent observers on the same reach represents a ‘significant difference’, implying an unsatisfactory result. The information on aquatic macrophyte cover variability is clearly of a fragmentary and unsatisfactory nature and needs further study. It would be surprising if aquatic plant cover estimates show as much precision as those of terrestrial plant estimates, given the working conditions. There are the added problems of water reflection, refraction, turbidity and accessibility. Another difficulty is the fragile nature of the vegetation. If several observers record from the same reach, losses are likely to result from uprooting, trampling and smothering by silt (Wiegleb, 1989). These factors may serve to amplify inter-observer variability without directly relating to observer effects.

3.1.2 Estimation of species richness

Macrophyte surveys usually aim to identify all species within a river reach, or all those species present on a standard recording sheet. The relevance of this practice depends on the use to which the information is to be put. In calculating some metrics, species are assigned ranks depending on their ecological preferences or optima (whether based on expert judgement or empirical data) and these ranks are then multiplied by a cover value to yield the required metric. If species are misidentified or overlooked, a large error in the final metric may occur, especially in species-poor reaches. Species considered to be more reliable indicators may be given extra weight, or, as in MTR, some species considered indifferent to a particular pressure may be discounted. In other cases survey data is used in a less refined form, with the focus on determining the presence of characteristic or type-specific species. However all surveys, no matter how comprehensive, can only provide a sample of the biota present. The ratio of the sample to the true community will vary between observers, between river types and with survey conditions and these factors are likely to also have interactive effects on survey efficiency. Simulation exercises (such as Ewald, 2003) indicate that derived metrics based on composition are relatively insensitive to variable detection rates, or to random deletion of taxa with low cover values. As a stand-alone metric, richness is likely to show much greater variability between observers than compositional metrics, although it will conceal differences in detection of individual species between observers. Comparisons of raw biological composition will be highly vulnerable to variable detection rates and misidentifications (Lansdown, 2007).

A form of quality control is usually introduced to attempt to reduce the error in recording by employing a minimum of two observers to survey the same reach. On the basis that two pairs of eyes are better than one, this should increase detection rates. Use of fixed recording lists is also known to reduce oversight in terrestrial surveys.

Species-richness is extremely variable in rivers, and there is a suggestion that it attains a maximum in mesotrophic waters (Dawson *et al*., 1999; Bornette *et al*., 2001). The range in richness recorded in reaches of 100-m is considerable. Springe & Sandin (2004) give a range of one to 20 species in 25 reaches, and Staniszewski *et al*. (2006) found 17-85 species in 43 reaches of 100-m. Coops (in Birk *et al*. 2007) found from one to 60 taxa in some 100-m Dutch river reaches. Baatrup-Pedersen *et al*. (2006) looked briefly at the relationship between species richness and area, recording 14 species for a particular type of reach. A jackknife method suggested the total number of species was in fact 23. (The jackknife is a statistical technique for obtaining estimates based upon the recomputation of a statistic from a sample, removing one observation at a time).

River width is usually recorded in macrophyte surveys, but is rarely reported in publications. As a result it is not possible to comment further on species-area relationships for aquatic macrophytes (although the data needed to derive such relationships could probably be obtained or estimated, via basic relationships between width and factors such as distance downstream). Since species richness in whole EU river systems is often in the range 70-140 (Holmes & Whitton, 1977; Bornette, 2001), single reach lists usually fall short of the total list for the river in question. They may, on occasion, fall short of the recommended area based upon species-area curves. Further investigation of this relationship has been suggested in BSI (2003). Determinations of sample area are not always straightforward, especially with regard to river bank position. For aquatic macrophytes this has been taken as the water level corresponding to permanent submersion for greater than either 50 or 85 per cent of the time. This may be difficult to estimate, especially where the bank has a low angle of slope. Consequently, some variability between observers in recording will reflect uncertainties in designation of the sampling area.

3.2 Estimation of derived metrics

While there is little published information on the variability of primary metrics, some data are available for derived metrics, particularly MTR, IBMR and some universal ecological metrics such as species diversity and evenness. It has been argued that it is only the variability of these derived metrics that is important, since they are the quantities that yield, more or less directly, a river classification.

Although there is no general agreement on the numerical definition of ‘precision’, the sample standard deviation is usually taken as an acceptable measure (APHA, 2004), but the half-width of the confidence interval has also been used (Ellis & Adriaenssens, 2006). However, confidence intervals vary according to the number of measurements taken. The standard deviation is an estimate of a population standard deviation that is usually unknown, but is a constant for a defined population of measurements. Sample standard deviations are obtained from a series of repeated measurements and ideally the number of replicates should be at least ten. The larger the number of measurements taken, the closer the sample standard deviation will be to the population constant. Standard deviations based on a few measurements, say two or three, are of little value and are not likely to provide a useful estimate of precision. Precision is often quoted as the coefficient of variation (CV), the standard deviation as a percentage of the mean. Typical and acceptable values of precision for chemical analyses are around two to five per cent. Standard errors can also be used to define precision, but only where there is sufficient replication.

Four published studies of metrics such as the MTR that include descriptive statistical data were found. Dawson *et al*. (1999) reported on data collected from the UK with sites assessed by two different surveyors at slightly different times in the same year, with the assumption that any variability would be principally due to inter-surveyor error. The two surveys (an initial 'primary' survey and a later 'audit' survey) were compared by examining the 41 differences in the two MTR scores. Presented in histogram form, these differences were considered approximately Gaussian with a standard deviation of 6.2 MTR units. They were also obtained as a percentage difference from the primary MTR scores. A difference measured in percentages would not be expected to conform to a normal distribution, but the standard deviation was 17.8 percentage units. These results can also be expressed in terms of frequency – 68 per cent of the audit surveys differed by less than 15 per cent from the primary surveys. While this is a useful approach to the problem, the data are insufficient for rigorous statistical testing and can give no more than an idea of the magnitude of inter-surveyor error. This is because there were no precision surveys of the sites against which the observer error could be tested. 'Audit' surveys would not qualify as they still have an unknown observer error, even though they would have been undertaken by experts.

Dawson *et al*. (1999) also looked at the effects of plant misidentification and cover mis-classification. They used a Monte-Carlo modelling approach and concluded that an error of cover value of one class makes little difference to the MTR, but an error in identification of a dominant, as defined by its Species Trophic Rank (STR) can make a difference of up to 7.5 MTR units. Assigning precision values to this source of error will be observer-dependent and thus difficult to determine. Particular problems were noted, namely the identification of Batrachian *Ranunculus* species and of selected aquatic bryophytes.

Springe & Sandin (2004) used a different approach to the study of variation in the MTR. They made a hierarchical investigation of three Estonian river basins with nests of three 100-m samples in each of three rivers giving a total of 27 samples, although only 25 appear in the analysis. Each site was sampled once. They provide means, standard errors and CVs for the nested samples (Table 3.1). Sample standard deviations were calculated from their CVs. The 2004 study was more of a hierarchical comparison than a precision exercise. It was undertaken between river sections, rivers within basins and between basins. It showed that differences between reaches of the same river were the most variable using both MTR and IBMR although it is not clear whether the standard deviation was calculated for all 25 samples, or was an average for the three reaches taken in each river. The study also showed that the aquatic vegetation of the three river basins, and indeed for all of the sites, gave remarkably similar scores using the above metrics. This was also demonstrated statistically.

Table 3.1 Data on the precision of some derived metrics from Springe & Sandin (2004) rearranged and modified.

Sample number	Type of sample		MTR			IBMR	
		Mean	Standard deviation	CV	Mean	Standard deviation	CV
25	River reach samples	41.9	7.45	17.8	10.7	1.46	13.6
9	3 rivers in 3 basins	41.5	4.10	9.9	10.6	1.07	10.1
3	3 basins	41.5	2.74	6.6	10.6	0.75	7.1

More useful studies in terms of precision are those of Szoszkiewicz *et al*. (2004) and Staniszewski *et al*. (2006). Both papers analyse macrophyte data from Poland as part of the STAR (Standardisation of River Classifications) Project. Many results are common to both but the paper by Szoszkiewicz *et al*. (2004) contains more information. In this study, 43 sites were selected to cover a wide trophic range and were spread through two ecoregions, XIV and XVI. Reaches of 100-m were sampled following the MTR protocol and a range of derived metrics were obtained including MTR and IBMR. Although 43 sites are mentioned in the publication, the tables provide data for only 26 sites where they are listed as 'streams'. It is unclear whether some of these 26 sites actually include more than one 100-m reach. It is assumed that only one reach is included.

This study looked at inter-surveyor variation on single 100-m reaches. Three surveyors independently assessed the same reach, although a total of six surveyors were used, and individuals were not tagged in the analysis. Table 3.2 shows the inter-surveyor, yearly and seasonal variation in four metrics, IBMR, MTR, and species richness. For each metric, the mean, standard deviation and CV are shown. It is important to note that the standard deviations and CVs are means of 26 sets of three (observer) or two (temporal) measurements. For example, the CV was first calculated for one set of three MTR values using three different observers on one reach. The process was then repeated for the remaining 25 sets of three. The average of the entire set of 26 standard deviations was then determined. This was done to obtain more reliable statistics. It would have been better to have used a larger number of replicates, but this would have been expensive.

The results show that the average differences between three observers are small. The 'averaged' CV of four to five per cent indicates good precision for the biological metrics MTR and IBMR and is similar in magnitude to the effect of variation between years and seasons. Estimates of species richness exhibited markedly higher variability due to surveyor, although this was much less than the variability associated with year and season. In addition Dawson *et al*. (1999) draw attention to increased observer experience of a site over time but they did not analyse year to year data. This means that the contribution of surveyor variability to total between-year variability (whether within or between surveyors) will not be constant over time.

Table 3.2 Precision of three factors relevant to macrophyte surveys in rivers Modified and partially recalculated from Staniszewksi *et al*. (2004) and Szoskiewicz (2006).

Factor	Replication		MTR			IBMR			RICHNESS	
		Mean	S.D.	CV	Mean	S.D.	CV	Mean	S.D.	CV
Surveyors	3	36.3	1.62	4.75	8.8	0.34	3.83	22.0	2.67	14.4
Year	2	37.6	2.47	6.89	9.4	0.36	3.83	22.8	5.09	24.5
Season	2	37.0	2.18	6.07	8.9	0.34	3.81	22.1	4.27	20.9

3.3 Derived metrics - temporal and longitudinal variation

Dawson *et al*. (1999) looked at within-year variation of MTR by comparing values obtained near the beginning and end of the summer growing season, a difference of about 110 days. They found that the MTR was, on average, 7.5 per cent higher at the end of the season compared with the beginning, with MTRs ranging from 10 to 74. They concluded that this was probably due to increasing plant cover values as the season progressed 'improving' the MTR score. However, cover values themselves should not greatly influence an MTR score. A related explanation may be that increased cover of high MTR scoring species over the course of the growing season increases the detectability of these species by surveyors and thereby elevates the MTR score. However, no comparisons of species richness over the duration of the survey period were made so it is impossible to confirm this explanation.

There is a greater problem in conducting experiments on natural systems when inter-year variation is to be examined, since a proportion of between-year variation may be related to genuine changes in the macrophyte community of the type which sampling is seeking to detect. In practice, within-year (between seasons) variability attributable to natural fluctuations in flow and water chemistry could be expected to be similar in magnitude to between-year variability. Directional changes in such variables over longer periods, possibly as a result of anthropogenic impacts, should be registered as increased between-year variability. Szoszkiewicz *et al*. (2004) and Staniszewski *et al*. (2006) looked at the effect of temporal variation at two levels, first on a yearly basis (sampling was undertaken in two years), and second on a seasonal basis. In one of the years, sampling was carried out in both summer and autumn. A similar pattern is apparent for the between-seasons data (summer/autumn) and the between-year data although the variability is higher in most cases (Table 3.2). It is evident that the seasonal variation is larger than the observer variation. This exercise shows that it can be inadvisable to treat the variation in derived metrics over different seasons of the same year as a surrogate for observer error.

If for the moment, a single site and single annual sampling is considered, how different do MTR values need to be to show change over that period? Reference to the above precision values is insufficient to decide over a single year. Dawson *et al*. (1999) on the basis of the statistical work reported above, suggest that a difference of 7.5 per cent in

MTR between sampling periods should be considered 'significant'. This, however, is not statistical significance. Even a large difference in the MTR value should be suspect and not taken in isolation (in fact, large differences are unlikely to occur between consecutive sites in a water body, or at the same site over the short term, and should be considered more dubious than small-medium scale differences). An unfortunate combination of highly contrasting weather, with extreme observer error over two sampling periods could lead to a difference much larger than 7.5 per cent but would bear no relationship to nutrient status. For example, in the data set of Szoszkiewicz *et al*. (2004), on one occasion two observers obtained MTR values of 18 and 30 for the same reach, a difference of 12 units or 60 per cent.

Longitudinal variation has received little direct investigation. Dawson *et al*. (1999) looked at the change in MTR down some moderately large UK rivers and found that in some, MTR values fell more or less regularly downstream. However, in others, there was little difference downstream. No reason was given and it was not thought that point-source pollution influenced any of the results. Wiegleb *et al*. (1989) found that more variability was encountered among sites than within sites sampled at different times. From a classification perspective, the change in absolute metric values along a water course is of less interest than the change in the ratio between observed and expected metric values (the EQR), since this indicates how the quality is changing relative to an unimpacted state, rather than simply how the biology is changing independently of some point of reference. This issue is relevant to the placement of sample sites since, given limited resources, there may be advantages in assigning sample sites to the downstream end of water bodies where the impacts of multiple, and possibly diffuse pressures upstream, are perhaps most likely to be recognisable.

Attempts have also been made to compartmentalise the variance between the several factors investigated although to date, sampling schemes have not allowed a comprehensive study of these. For instance, Ellis & Adriaenssens (2006) recognise the potential significance of spatio-temporal interaction but it remains to be tested. Johnson *et al*. (2006) used a Canonical Correspondence Analysis (CCA) to partition variability of macrophytes in some lowland and upland European rivers. They found that the geographic region (ecoregion) accounted for 2.8 per cent of the total variation and latitude 1.9 per cent. More variance was accounted for by some of the chemical determinands such as conductivity. In the montane rivers, higher percentages were attributed to the above categories than in the lowlands. As might be expected, Baatrup-Pedersen *et al*. (2006) found that macrophyte variability was higher within pre-defined typologies and less within botanical classifications made using TWINSPAN.

3.4 Quality control

Training schemes are in place for IBMR and MTR to ensure an acceptable standard of sampling and recording of macrophytes. The use of a minimum of two recorders per reach has been advocated for MTR and other schemes such as the GRI (German Reference Index) (see Holmes *et al*., 1999). This improves the efficiency of the survey process and individual measurements permit cross-referencing of qualitative and quantitative data and should be promoted. An occasional 500-m survey to check against a 100-m survey has also been mooted (Holmes *et al*., 1999). In addition, random samples are sent to a specialist for confirmation of identity from time to time. There is little information on other methods, but since many are based on MTR they probably follow a similar protocol.

Opportunities exist for a more formal audit scheme where a proportion of sites are revisited in the same season by an experienced surveyor. Criteria to define acceptable levels of similarity between surveys must be realistic, should take account of potential changes that have occurred in the intervening period, and should focus on the detection and correct identification of those taxa that characterise a reach, rather than on comparatively trivial issues such as similarities in cover values.

4 Conclusions and recommendations

1. While much has been achieved in the design and implementation of methods for assessing aquatic macrophytes in EU rivers, more work on the different sources of variability within the methods and sampling sizes is required.

2. There is currently good information on methods of aquatic macrophyte sampling used within Europe.

3. The timing of surveys is confined to the summer growth period throughout the EU countries surveyed and ranges from three to five months duration depending on local climate. In most cases, one or two samples are taken from sites every three years.

4. A sampling length of 100-m is used widely across EU states and the choice of length is justified more on practical than theoretical grounds. There are advantages in using 100-m in terms of convenience: for example, the mosaic-like pattern of lowland European land use and drainage density patterns in Northern Europe. There are disadvantages in large (greater than 50-m wide) rivers and small montane rivers due to the contrasting areas being sampled and physical nature of montane rivers which varies considerably over small spatial scales.

5. For the primary underived metrics such as species cover and species richness, the few studies that have been made on observer error suggest that misidentification and the overlooking of taxa are more significant than errors in estimation of cover, but a critical assessment cannot be made on the basis of the available published data.

6. For the derived metrics, such as MTR and IBMR, errors in estimating cover are usually less important than errors in identification or the overlooking of species. Since species are assigned different ranks, errors are difficult to determine as they depend on the ranks of overlooked or misidentified taxa. Observer precision for the MTR appears to be about five per cent (as coefficient of variation). Variability in estimates of species richness is much more variable between observers.

7. Within-season variation in MTR has not been studied widely, but a variation of up to ± 7.5 per cent has been reported and attributed to change in plant growth patterns as the season progresses. Another study estimates a coefficient of variation of around six per cent.

8. Between-year variation shows higher variability than between-season or observer error for metrics the IBMR and MTR. A coefficient of variation of about seven per cent has been reported.

The following recommendations are made to improve our understanding of macrophyte monitoring and ecology.

1. A better understanding of the relationship between species richness and length or area is needed, to help determine the best reach length to use. It is likely that such studies will suggest different reach lengths for different types or sections of river and a compromise may be required. Such a study could be usefully combined with a fractal analysis using semivariograms. Although decisions on reach length are likely to be standardised across EU countries, the above study could provide justification for the reach length(s) finally decided upon.

2. Assessments of the degree of species turnover between standard samples in different river types could be used to guide decisions of how many samples are required to obtain a representative sample of the vegetation.

3. Observer error has received some attention but would benefit from further study. This applies particularly to errors due to misidentification and the under-detection of species. Such a study would need to be designed with care and with provision for a better understanding of errors in cover estimation. Information on specific components of observer-based variability could assist in the refinement of training schemes and indicate suitable criteria for auditing the quality of surveys.

4. To further investigate spatio-temporal variations in macrophyte communities, a hierarchical study within river basin(s) with good replication is strongly recommended to clarify the variability between reaches, seasons and years.

References

AMERICAN PUBLIC HEALTH ASSOCIATION (APHA), 2004. Standard Methods for the Examination of Water and Wastewater.

BAATRUP-PEDERSEN, A., LARSEN, S.E. AND RISS, T., 2002. Long-term effects of stream management on plant communities in two lowland streams. *Hydrobiologia*, 481, 33-45.

BAATRUP-PEDERSEN, A., SZOSZKIEWICZ, K., NIJBOER, R., O'HARE, M. AND FERREIRA, T., 2006. Macrophyte communities in unimpacted European streams: variability in assemblage patterns, abundance and diversity. *Hydrobiologia*, 566, 179-196.

BARENDREGT, A AND BIO AMF, 2003. Relevant variables to predict macrophyte communities in running waters. *Ecological Modelling*, 160, 205-217.

BIRK, S., KORTE, T. AND HERING, D., 2006. Intercalibration of assessment methods for macrophytes in lowland streams: direct comparison and analysis of common metrics. *Hydrobiologia*, 566, 417-430.

BIRK, S., WILLBY, N., CHAUVIN, C., COOPS, H.C., DENYS, L., GALOUX, D., KOLADA, A., PALL, K., PARDO, I., POT, R. AND STERLZER, D., 2007. Central Baltic River GIG Macrophyte Intercalibration Exercise, June 2007. 82pp.

Available from: http://starwp3.eu-star.at

BORNETTE, G., PIEGAY, H., CITTERIO, A., AMOROS, C., GODREAU, V., 2001. Aquatic diversity in four river flood plains: a comparison at two hierarchical levels. *Biodiversity and Conservation*, 10, 1683-1701.

BSI, 2003. Water quality guidance standard for the surveying of aquatic macrophytes in running waters. British Standard BS EN 14184:2003. British Standards Institute.

BUTCHER, R.W., 1933. Studies on the ecology of rivers. I. On the distribution of macrophytic vegetation in the rivers of Britain. *Journal of Ecology*, 21, 58-91

CLARKE, P.T., 2004. STARBUGS 1.1 (Star Bioassessment Uncertainty Guidance software). Error/uncertainty software. CEH Dorset.

DAWSON, F.H., NEWMAN, J.R., GRAVELLE, M.J., RAVEN, K.J. AND HENVILLE, P., 1999. Assessment of the Trophic Status of Rivers using Macrophytes. Bristol: Environment Agency R & D Technical Report E39. 130pp.

ENVIRONMENT AGENCY, Operational Guidance Manual, 2007. Surveying freshwater macrophytes in rivers [draft report in progress of completion].

ELLIS, J.A. AND ADRIAENSSENS, V., 2006. Uncertainty estimations for monitoring results by the WFD biological classification tools. Environment Agency WFD Report. 32pp.

EWALD J., 2003. The sensitivity of Ellenberg indicator values to the completeness of vegetation releves. *Basic and Applied Ecology*, 4, 507-513.

FRISSEL, C., LISS, W.J., WARREN, C.E. AND HURLEY, M.D., 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental Management*, 10, 199-214.

GREIG-SMITH, P., 1957. Quantitative Plant Ecology. London: Butterworth. 198pp.

HAURY, J. AND MULLER, S., 1991. Variations écologiques et chorologiques de la végétation macrophytique des rivières acides au Massif Armorican et des Vosges du Nord (France). *Revue des Sciences de l'Eau*, 4, 463-482.

HAURY, J., PEUTRE, M-C., TREMOLIERES, M. AND BARBE, J., 2002. Proceedings of the 11th EWRS International Symposium on Aquatic Weeds, 247-250.

HOLMES, N.T.H., 1980. Preliminary results from river macrophyte survey and implications for conservation. Nature Conservancy Council, Chief Scientists Team, Note 24.

HOLMES, N.T.H., 1991. British river macrophytes – perception and use in the 20th century. *Aquatic Conservation of Marine and Freshwater Ecosystems*, 9, 535-539.

HOLMES, N.T.H., NEWMAN, J.R., CHADD, S., ROUEN, K.J., SAINT, L. AND DAWSON, F.H., 1999. Mean Trophic Rank: A User's Manual. Environment Agency R & D Technical Report E38. 132 pp.

HOLMES, N.T.H. AND WHITTON, B.A., 1977. The macrophyte vegetation of the River Tees in 1975: observations and predicted changes. *Freshwater Biology*, 7, 43-60.

JNCC, 2005. Notes on Common Standards for Rivers. Joint Nature Conservation Committee.ISSN 1743-8160. 58pp.

JOHNSON, R.K., FURSE, M.T., HERING, D. AND SANDIN, L., 2007. Ecological relationships between stream communities and spatial scale: implications for designing catchment-level monitoring programmes. *Freshwater Biology*, 52, 939-958.

KERSHAW, K. A. AND LOONEY, J.H.H., 1985. Quantitative and Dynamic Plant Ecology. Hodder Arnold.

LACOUL, P. AND FREEDMAN, B., 2006. Environmental influences on aquatic plants in freshwater ecosystems. *Environmental Reviews*, 14, 89-136.

LANSDOWN, R.V., 2007. Assessment of the condition of the water courses of plain to montane levels with the Ranunculion fluitantis and Callitricho-Batrachion vegetation of the River Wye SAC. Countryside Council for Wales.

O'HARE, M. T., BAATRUP-PEDERSEN, R., NIJBOER, K., SZOSZKIEWICZ, K. AND FERREIRA, T., 2006. Macrophyte communities of European streams with altered physical habitat. *Hydrobiologia*, 566, 197-210.

PALMER, M.W., 1988. Fractal geometry: a tool for describing spatial patterns of plant communities. *Vegetatio*, 25, 91-102.

PASSAUER, B., MELINGER, P., MELZER, A. AND SCHNEIDER, S., 2002. Does the structural quality of running waters affect the occurrence of macrophytes? *Acta Hydrochimica et Hydrobiologia*, 30, 197-206.

RICHARDS, K.S., 1976. Channel width and riffle-pool sequence. *Geological Society of America*, 87, 883-890.

RUXTON, G.D. AND COLEGRAVE, N., 2003. Experimental Design for the Life Sciences. Oxford University Press.

SCHAUMBERG, J., SCHRANZ, C., FOERSTER, J., GUTOWSKI, A., HOFMANN, G., MEILINGER, P., SCHNEIDER, S. AND SCHMEDTJE, U., 2004. Ecological classification of macrophytes and phytobenthos in Germany according to the Water Framework Directive. Available from: http://www.bayern.de/lfw/projeckte/welcome.htm

SIRJOLA, E., 1969. Aquatic vegetation of the River Teuronjoki, south Finland, and its relation to water velocity. *Annales Botanici Fennici*, 6, 68-75.

SPRINGE, G. AND SANDIN, L., 2004. Calibrating bioliogical survey results.Objective to estimate different types of variability in macrophyte assessment.

Available from: http://www.eu-star.at/pdf/Deliverable_N3.pdf

STANISZEWSKI, R., SZOSZKIEWICZ, K., ZBIERSKA, J., LESNY, J., JUSIK, S. AND CLARKE, R.T., 2006. Assessment of source of uncertainty in macrophyte surveys and the consequences for river classification. *Hydrobiologia*, 566, 235-246.

SZOSZKIEWICZ, K., ZBIERSKA, J., STANISZEWSKI, R., JUSIK, S., ZGOLA, T. AND KUPIEC, J., 2004. Errors and variation associated with field protocols for the collection and application of macrophyte and hydro-morphological data.

Available from: http://www.eu-star.at/pdf/Deliverable_N4.pdf

WHITTON, B.A. AND BUCKMASTER, R.C., 1970. Macrophytes of the River Wear. *Naturalist*, 97-116.

WILKINSON, J., MARTIN, J., BOON, P. AND HOLMES, N.T.H., 1998. Convergence of field survey protocols for SERCON (System for Evaluating Rivers for Conservation) and RHS (River Habitat Survey). *Aquatic Conservation Marine and Freshwater Ecosystems*, 8, 579-596.

WIEGLEB, G., HERR, W. AND TODESKINO, D., 1989. Ten years of vegetation dynamics in two rivulets in Lower Saxony (FRG). *Vegetatio*, 82, 163-178.

WILLBY, N.J., PITT, J.A. AND PHILLIPS, G. 2009. Development of a system for the classification of lakes and rivers in the UK using aquatic macrophytes. Part II. Rivers. Environment Agency Science Report.

UKTAG, 2007. UK Classification Scheme for Surface Waters. UK Technical Advisory Group on the Water Framework Directive. Draft report, September 2007

List of abbreviations

CCW Countryside Commission for Wales

CV Coefficient of variation

CSM Common standards monitoring

GRI German Reference Index

IBMR Indice Biologique Macrophytique Rivière

JNCC Joint Nature Conservation Committee

MTR Mean Trophic Rank

RHS River Habitat Survey

SEPA Scottish Environmental Protection Agency

SNH Scottish Natural Heritage

SSSI Site of Special Scientific Interest

STAR Standardization of River Classifications

Glossary

Derived metric	Ecological measurement obtained by combining two or more underived metrics.
Underived metric	Ecological measurement such as percent cover of vegetation or pH.

Would you like to find out more about us, or about your environment?

Then call us on

08708 506 506* (Mon-Fri 8-6)

email

enquiries@environment-agency.gov.uk

or visit our website

www.environment-agency.gov.uk

incident hotline 0800 80 70 60 (24hrs)
floodline 0845 988 1188

* Approximate call costs: 8p plus 6p per minute (standard landline).
Please note charges will vary across telephone providers

Environment first: This publication is printed on recycled paper.