

MONITORING BIOLOGICAL INDICATORS FOR THE WFD IN SWEDISH WATER BODIES

Current designs and practical solutions for quantifying overall uncertainty and its components

Mats Lindegarth, Jacob Carstensen, Richard K. Johnson

WATERS Report no. 2013:6

WATERS Report no. 2013:6
Deliverable 2.2-2

Monitoring biological indicators for the WFD in Swedish water bodies

Current designs and practical solutions for quantifying
overall uncertainty and its components

Mats Lindegarth, University of Gothenburg

Jacob Carstensen, Aarhus University

Richard K. Johnson, Swedish University of Agricultural Sciences

WATERS partners:



WATERS: Waterbody Assessment Tools for Ecological Reference conditions and status in Sweden

WATERS Report no. 2013:6. Deliverable 2.2-2

Title: Monitoring biological indicators for the WFD in Swedish water bodies: Current designs and practical solutions for quantifying overall uncertainty and its components

Publisher: Havsmiljöinstitutet/Swedish Institute for the Marine Environment,
P.O. Box 260, SE-405 30 Göteborg, Sweden

Published: October 2013

ISBN 978-91-9806646-9-8

Please cite document as:

Lindegarh, M., Carstensen, J., Johnson, R.K. Monitoring biological indicators for the WFD in Swedish water bodies: Current designs and practical solutions for quantifying overall uncertainty and its components. Deliverable 2.2-2, WATERS Report no. 2013:6. Havsmiljöinstitutet, Sweden.

<http://www.waters.gu.se/rapporter>

WATERS is a five-year research programme that started in spring 2011. The programme's objective is to develop and improve the assessment criteria used to classify the status of Swedish coastal and inland waters in accordance with the EC Water Framework Directive (WFD). WATERS research focuses on the biological quality elements used in WFD water quality assessments: i.e. macrophytes, benthic invertebrates, phytoplankton and fish; in streams, benthic diatoms are also considered. The research programme will also refine the criteria used for integrated assessments of ecological water status.

This report is a deliverable of one of the scientific sub-projects of WATERS dealing with uncertainty of current monitoring programmes in the perspective of the EU Water Framework Directive. We analyse sources of uncertainty arising from the structure of monitoring, identify components of uncertainty that might need further attention and finally to suggest methods statistical and empirical methods for quantifying these components. These results will guide further work within WATERS and provide input to on-going efforts to improve national and regional monitoring.

WATERS is funded by the Swedish Environmental Protection Agency and coordinated by the Swedish Institute for the Marine Environment. WATERS stands for 'Waterbody Assessment Tools for Ecological Reference Conditions and Status in Sweden'. Programme details can be found at: <http://www.waters.gu.se>

Table of contents

Executive summary	9
Svensk sammanfattning	10
1 Introduction	11
1.1 Monitoring according to the WFD	12
1.2 Uncertainty in the WFD and in Swedish assessment criteria	13
1.3 The uncertainty framework	15
2 Objective	17
3 Structure, dimensioning, and uncertainties of current monitoring designs	18
3.1 Conceptual analysis of alternative sampling designs	18
3.1.1 Crossed designs in a single water body	18
3.1.2 Nested design in a single water body	20
3.1.3 Crossed and nested designs in a water body type	22
3.2 Structure of current monitoring programmes	23
3.2.1 Coastal waters	24
3.2.2 Lakes	26
3.2.3 Streams	29
3.2.4 General features and conclusions	30
4 Sampling designs to quantify uncertainty components.....	32
4.1 Principles of sampling designs	33
4.2 Quantifying uncertainties not estimated by monitoring programmes.....	34
4.2.1 Estimation by reduced designs.....	34
4.2.2 Estimation using information from other areas or times	37
4.2.3 Data requirements for the estimation of uncertainty components	38
5 Conclusions	41
6 Acknowledgements.....	42
7 References	43

Executive summary

This report uses a comprehensive uncertainty framework (Lindegarth et al. 2013) for analysing and reviewing the monitoring requirements for biological quality elements (BQEs), as defined in the WFD, and for the general spatial and temporal structure of existing monitoring in Swedish coastal and inland waters. The study aims to 1) examine the complexity of potentially important sources of uncertainty arising from the monitoring structure for particular BQEs, 2) to identify components of uncertainty that might need further attention and 3) to suggest statistical and empirical methods for quantifying these components.

The general conclusion is that the framework provides a useful tool for analysing uncertainties in a wide range of situations, and the analyses identify a number of general and specific properties of current monitoring that need to be addressed to ensure appropriate assessment and, ultimately, reduced uncertainty. Furthermore, we identified a large variety of monitoring approaches and therefore also large differences in the combination of relevant uncertainty components. It was noted that, in assessing individual water bodies, there is frequently a lack of replicate sites (or stations), potentially causing a lack of spatial representativity. However, it is concluded that this lack of replication at water body scale may not cause severe problems at the water body type or catchment scales, because assessment of status and uncertainty at these scales could be representative due to replication in a number of water bodies. This insight is of particular relevance for coming status assessments according to the MSFD.

We also illustrated how the uncertainty framework can be used in combination with existing data or the strategic addition of replicates at selected spatial scales to quantify critical components of uncertainty. For this purpose, we presented alternative designs, i.e., nested or staggered sampling designs, and illustrated methods for assessing the expected precision of estimates of variance components. For example, these methods indicate that the average deviation of an estimate from its true value (i.e., the standard error of the estimated variance) is 20–25% when the variance is estimated with 30–50 degrees of freedom. Although any rules of thumb for how precise a variance estimate needs to be are somewhat arbitrary, they do provide useful tools for WATERS' coming estimation of uncertainty components, for the compilation of an “uncertainty library” (suggested by Lindegarth et al. 2013), and for any authority responsible for assessing uncertainty in the classification of ecological status.

Svensk sammanfattning

I denna rapport analyseras kraven på miljöövervakning av biologiska kvalitetsfaktorer enligt EU:s vattendirektiv, och den rumsliga och tidsmässiga strukturen hos pågående svensk övervakning. Detta görs från perspektivet av den heltäckande metod för hantering av mätosäkerhet som utvecklades av Lindegarth et al. (2013). Syftet med studien är att 1) illustrera hur övervakningens utformning på ett komplext sätt påverkar osäkerheten i statusklassning av olika biologiska kvalitetsfaktorer, 2) att identifiera osäkerhetskomponenter som kan behöva ytterligare uppmärksamhet och att 3) föreslå empiriska och statistiska metoder för att bestämma storleken på dessa komponenter.

Den övergripande slutsatsen är att den föreslagna metodiken på ett bra sätt kan användas för att analysera osäkerhet i ett brett spektrum av situationer. Dessa analyser identifierar en rad generella och specifika egenskaper hos strukturen av dagens övervakning, som behöver hanteras för att åstadkomma en tillfredsställande övervakning och för att i slutändan minska osäkerheten. Dessutom identifierades stora olikheter mellan kvalitetsfaktorer, i övervakningens utformning och följaktligen även skillnader i vilka osäkerhetskomponenter som kan tänkas vara betydelsefulla. En generell slutsats är att övervakningen av enskilda vattenförekomster ofta baseras på enskilda provtagningslokaler (stationer) vilket kan leda till brist på rumslig representativitet. Det noteras dock att denna brist inte behöver innebära några problem på nivån vattentyper eller avrinningsområden eftersom representativa skattningar av status och osäkerhet på dessa nivåer kan åstadkommas genom övervakning i flera vattenförekomster. Denna insikt är av särskild betydelse för kommande statusbedömningar enligt EU:s havsmiljödirektiv.

Vi illustrerar också hur metoden för osäkerhetshantering kan användas med befintliga data eller genom att man på ett strategiskt sätt kompletterar befintlig provtagning på lämpliga rumsliga skalor, för att mäta viktiga osäkerhetskomponenter. För detta ändamål beskriver vi på ett generellt plan möjliga alternativa provtagningsstrategier (på engelska benämnda "nested" och "staggered designs") och illustrerar metoder för att bedöma förväntad precision vid skattning av viktiga osäkerhetskomponenter. Exempelvis visar dessa metoder att det förväntade felet vid skattning av varianskomponenter (d.v.s. "standard error" för en skattad varians) är 20–25% när denna skattas med 30–50 frihetsgrader. Även om sådana tumregler för vad som är en lämplig precision i viss mån är godtyckliga kan metoderna användas för WATERS' planerade skattning av osäkerhetskomponenter och sammanställning av "osäkerhetsbibliotek" (Lindegarth et al. 2013) och vid behov inom myndigheter med ansvar för bedömning av osäkerhet vid klassning av ekologisk status.

1 Introduction

The Water Framework Directive (WFD) (2000/60/EC) was formulated to deal with the increasing pressures on European water resources and to achieve “good ecological status” in all European surface waters and groundwaters by 2015. The Directive defines a cyclic adaptive process, a number of administrative regulations, and several more or less specific guidelines for how the member states should implement the Directive in their respective countries and laws (Table 1.1).

Several components of this process specifically require that ecological data be collected in well-designed and appropriate monitoring programmes of various types. For example, monitoring ecological status and change over time is crucial to the characterisation of a water body, evaluating environmental objectives and assessing the efficiency of management plans. Therefore, designing and implementing cost-effective and precise monitoring programmes is a fundamental requirement for meeting the objectives of the Directive. General principles for monitoring under the WFD are outlined in the CIS Guidelines #7 (EC 2003).

TABLE 1.1

Components, timetable, and references to WFD definitions. Modified from http://ec.europa.eu/environment/water/water-framework/info/timetable_en.htm.

Year	Issue	Reference
2000	Directive entered into force	Art. 25
2003	Transpose Directive into national legislation	Art. 23
	Identify River Basin Districts and Authorities	Art. 3
2004	Characterise river basins: pressures, impacts, and economic analysis	Art. 5
2006	Establish monitoring network	Art. 8
	Start public consultation (at the latest)	Art. 14
2008	Present draft river basin management plan	Art. 13
2009	Finalise river basin management plan, including programme of measures	Art. 13 & 11
2010	Introduce pricing policies	Art. 9
2012	Make operational programmes of measures	Art. 11
2015	Meet environmental objectives	Art. 4
	First management cycle ends	
	Second river basin management plan and first flood risk management plan	
2021	Second management cycle ends	Art. 4 & 13
2027	Third management cycle ends, final deadline for meeting objectives	Art. 4 & 13

1.1 Monitoring according to the WFD

To support the requirements of the WFD cycle, the Directive outlines the need for three types of monitoring: surveillance, operational, and investigative monitoring (Table 1.2).

In terms of biological monitoring, surveillance monitoring is necessary for assessing the overall state and long-term change of the catchment (or subcatchment) in a River Basin District by monitoring a subset of water bodies using all biological quality elements (BQEs): phytoplankton, macrophytes, benthic invertebrates, and fish. The objective of surveillance monitoring is to assess water body status and long-term changes, for use in the design of future monitoring programmes and in conjunction with various types of impact assessments. In contrast, operational monitoring and investigative monitoring do not require that all BQEs be monitored, but instead the measurement of “quality elements which are indicative of the pressures to which the body or bodies are subject”.

Furthermore, operational monitoring targets water bodies identified as at risk of not meeting environmental objectives. Investigative monitoring seeks to understand why a particular water body does not meet its environmental objectives of “good” or “high” status.

TABLE 1.2

Objectives and guidelines for three types of monitoring in the WFD. Modified from 2000/60/EC, pp. 53–56.

Type of monitoring	Objective	Selection of monitoring points	Selection of quality elements
Surveillance	a. supplement and validate impact assessment b. design future monitoring programmes c. assess long-term changes in natural conditions d. assess long-term changes resulting from widespread anthropogenic activity	Insufficient surface water bodies to allow the assessment of the overall surface water status in each catchment or subcatchment in the River Basin District	a. parameters indicative of all BQEs b. parameters indicative of all hydromorphological quality elements c. parameters indicative of all general physico–chemical quality elements d. priority-list pollutants discharged into the river basin or sub-basin e. other pollutants discharged in significant quantities in the river basin or sub-basin
Operational	a. establish the status of bodies identified as at risk of failing to meet their environmental objectives b. assess any changes in the status of such bodies resulting from the programmes of measures	In all bodies of water that, based on either the impact assessment or surveillance monitoring, are identified as at risk of failing to meet their environmental objectives	Quality elements indicative of the pressures to which the water body or bodies are subject
Investigative	To be carried out when the reason for any exceedance is unknown and to ascertain the magnitude and impacts of accidental pollution	Not specified	Not specified

1.2 Uncertainty in the WFD and in Swedish assessment criteria

Although monitoring is usually the most reliable and objective method for obtaining information about the status of a water body or water body type, it is also true that all information derived from monitoring data is associated with errors and uncertainties. In a previous review, Lindegarth et al. (2013) outlined the fundamental principles for assessing

uncertainty defined by the WFD and their implementation in Swedish assessment criteria and legislation.

In short, the Directive and its guidance documents identify *precision* and *confidence* as the central concepts of uncertainty. Precision refers to the uncertainty of an estimated parameter (usually the mean), while confidence is a measure of the confidence associated with a certain classification, as in “the probability of the status being good or high is 75%”. Precision and confidence are determined by the data variability (s^2), the number of samples (n), and the desired level of confidence (i.e., risk of type 1 error, α). The confidence of a status assessment is also influenced by the differences (L) between the estimated mean and the class boundaries. Consequently, at a conceptual level, the fundamental principles for assessing uncertainty are well-defined and based on sound statistical principles. Nevertheless, the review also noted issues pertaining to the choice of level of acceptable confidence and decision rules (e.g., “face-value”, “fail-safe”, and “benefit-of-doubt”) that were not defined in the Directive. In conclusion, uncertainty in estimating and classifying biological indicators is clearly unavoidable in WFD status assessment procedures. The Directive and its guidance documents acknowledge this, provide useful definitions of uncertainty, and stipulate requirements for reporting, specifying that all assessments should be associated with estimates of the precision of and confidence in classification.

In Sweden, the WFD is implemented by chapter 5 of the Environmental Code, the Ordinance on Water Quality Management (Vattenförvaltningsförordningen, SFS 2004:660), and regulations from the Environmental Protection Agency (Naturvårdsverkets föreskrifter och allmänna råd om klassificering och miljökvalitetsnormer avseende ytvatten; NFS 2008:1). Guidance and advice on how to handle uncertainty are provided in the handbook *Status, potential och kvalitetskrav för sjöar, vattendrag, kustvatten och vatten i övergångszon* (2007:4) (see SEPA 2010 for a version in English). Lindegarth et al. (2013) concluded that: (1) the current Swedish assessment criteria do not cover all aspects of uncertainty as defined by the CIS guidance documents; (2) there are substantial conceptual differences among BQEs in how uncertainty is expressed and addressed in the assessment procedure; (3) for none of the BQEs does comprehensive guidance exist on how to handle various sources of uncertainty (e.g., spatial, temporal, and methodological). In particular, routines for addressing uncertainty at appropriate spatial and temporal scales (i.e., in a water body throughout a six-year assessment period) are currently lacking for all BQEs.

To address these deficiencies, Lindegarth et al. (2013) proposed a general framework that allows a more coherent and realistic estimation of precision and confidence than is permitted by current procedures. This framework can be used to analyse current monitoring designs and provides a solid foundation for attempts to reduce uncertainty by optimising monitoring designs and incorporating important environmental factors as covariates.

1.3 The uncertainty framework

Lindgarth et al. (2013) proposed that uncertainty should be assessed in the Swedish assessment WFD criteria by means of framework-based estimation of variance components using mixed models (e.g., Bolker et al. 2009). The framework applies general procedures for uncertainty (or error) propagation (e.g., Cochran 1977, Taylor 1997) and is based on scientific studies demonstrating the need for the combined assessment of various sources of uncertainty (e.g., Clarke et al. 2002, 2006a,b, Clarke & Hering 2006, Bennet et al. 2011, Mascaró et al. 2012). By explicitly adapting to temporal and spatial scales relevant to the WFD, the framework constitutes a general basis for further work in WATERS and in Swedish water quality assessment. It is also worth noting that a similar approach was used by Wikner et al. (2008) in developing a strategy for surveillance monitoring in the Bottenviken Water District.

The framework involves specifying a general linear model including random (CAPITAL letters) and fixed (lowercase letters) factors and interactions. These components can be categorised as temporal, spatial, and spatio-temporal interactions and variability associated with sampling and measurement:

$$\begin{aligned}
 y = \mu + & \underbrace{year + YEAR + season + SEASON \times YEAR + DIURNAL + IRREGULAR}_{\text{temporal sources of uncertainty}} \\
 & + \underbrace{gradient + GRADIENT + PATCHINESS}_{\text{spatial sources of uncertainty}} \\
 & + \underbrace{YEAR \times GRADIENT + SEASON \times GRADIENT}_{\text{spatio-temporal interactions}} \\
 & + \underbrace{sampling\ devices + PERSON + instrument + REPLICATE}_{\text{sampling and measurement uncertainties}}
 \end{aligned}$$

By estimating the size of these components, the variance ($V[\bar{y}]$) associated with a certain mean estimate (\bar{y}) can be estimated and partitioned into different sources of variability. Such partitioning is fundamental to the appropriate assessment of precision and confidence in classification and to the cost-benefit optimisation of monitoring programmes, both of which are necessary for the future development of water quality assessment routines according to the WFD. We present estimation procedures in their most basic form, i.e., when the components of variability, including residual deviations, are approximately normally distributed. Inspecting the residuals and testing these assumptions are recommended and, in cases of significant deviations, use of transformations or alternative link functions may be considered.

A general formulation of the total variance ($V[\bar{y}]$) affected by three random sources of variation (i.e., A, B, and C), each with a , b , and c levels, is that the sampling variance of a mean (\bar{y}) consists of three variance components, i.e., s_A^2 , s_B^2 , and s_C^2 . The combined total variance of the estimated mean, \bar{y} , is estimated from the size of the variance components and the number of levels:

$$V[\bar{y}] = \frac{s_A^2}{a} + \frac{s_B^2}{b} + \frac{s_C^2}{c}$$

To represent uncertainty, however, the total variability, $V[\bar{y}]$, needs to be transformed into a measure of the standard error of the mean, $SE_{\bar{y}} = \sqrt{V[\bar{y}]}$, and finally into a confidence interval according to

$$CI\% = [\sqrt{V[\bar{y}]} * t_{\alpha/2,df}; \sqrt{V[\bar{y}]} * t_{1-\alpha/2,df}]$$

where $t_{\alpha/2,df}$ and $t_{1-\alpha/2,df}$ are the percentiles of the t -distribution (usually the 2.5 and 97.5 percentiles, corresponding to $\alpha = 5\%$) with df effective degrees of freedom. If the degrees of freedom for $V[\bar{y}]$ exceed 30, the percentiles of the t -distribution can be approximated using the standard normal deviates, i.e., $z_{\alpha/2}$ and $z_{1-\alpha/2}$.

Using this general formula, the main priority in developing the uncertainty framework was to focus on assessment procedures at temporal and spatial scales relevant to the WFD. The main aim of the surveillance and operational monitoring in the WFD cycle is to assess water body status over six-year periods. This often implies that data from several sites and multiple years need to be combined and that the uncertainty of the estimated mean needs to be estimated. This is in contrast to the existing assessment criteria, which generally provide very little guidance on how to combine data from multiple years and no guidance on how to calculate the associated uncertainty. Furthermore, the surveillance monitoring should provide data from “sufficient surface water bodies to provide an assessment of the overall surface water status within each catchment”. This means that assessing overall uncertainty at the scales of catchments or water body types over six-year periods is also a high priority (in the marine environment, this is particularly relevant because in Sweden the Marine Strategy Framework Directive, MSFD, assessments will be conducted at the water body type scale). This outlines the main components of the uncertainty framework; for more details, see Lindegarh et al. (2013).

2 Objective

The objective of this report is to apply the proposed uncertainty framework to the context of current and future Swedish monitoring programmes aimed at fulfilling the requirements of the WFD. The overall aim is to analyse potential sources of uncertainty in current monitoring programmes, identify critical components, and suggest principles and designs that can be used by the authorities to quantify the importance of various uncertainty components.

After initial reviews of the monitoring guidelines set by the WFD and of the principles underlying the uncertainty framework, we conceptually illustrate the consequences of alternative designs for monitoring. We particularly focus on the spatial and temporal structure of sampling in individual water bodies and at larger scales of spatial aggregation (i.e., water body types or water catchments). Introducing these concepts allows us to analyse information on the structure of current monitoring in Sweden. Relevant information was extracted from the Swedish national database VattenInformationsSystem Sverige (VISS).

These analyses will help us identify uncertainty components that are poorly represented in current monitoring programmes but that are critical for the reliability of status assessments. With these components in mind, we present both theoretical and practical (i.e., cost-effective) sampling designs and analyses that can be used to quantify these critical components. Ultimately, this information can be used to (1) assess the uncertainty of status classifications using current and future monitoring designs and (2) to reduce uncertainty by providing guidelines for modifying monitoring designs.

3 Structure, dimensioning, and uncertainties of current monitoring designs

As demonstrated in earlier sections, the uncertainty of a status assessment depends on (1) the sampling variability of the particular indicator of interest, which is largely determined by biological spatio–temporal patterns, and (2) the structure and dimensioning of the sampling design, which are usually defined based on financial, practical, and historical constraints. To understand the uncertainties of current WFD assessments and ultimately to develop more reliable assessments, we review the structural properties of and conceptual issues concerning current Swedish designs for BQE monitoring. Together with quantitative estimates of sampling variability, these analyses will provide a basis for more efficient use of existing data and for optimizing future monitoring designs.

3.1 Conceptual analysis of alternative sampling designs

The WFD defines the spatial and temporal units for which ecological status needs to be assessed. These fundamental units are the water body and the six-year assessment period. To assess these in a spatially and temporally representative way, data can be collected in many ways. Data are usually collected at spatially discrete stations and temporally discrete times (often unevenly distributed among the six years). In terms of the spatial and temporal sampling structure, there are two principal strategies, which will be explained in detail below: (1) stations and times are sampled according to a crossed (also called orthogonal) design or (2) stations are sampled within time periods according to a nested design. For more complex spatio–temporal sampling involving more than two factors, combinations of crossed and nested designs are possible.

3.1.1 Crossed designs in a single water body

One monitoring design representative of most current programmes in aquatic environments in Sweden is one in which the same sites (“stations”) are revisited and sampled repeatedly year after year (Figure 3.1). The sites may have been selected completely at random in the water body or using criteria such as a narrow depth range, substrate, or distance from shore. The important thing is that the sites are often selected to “represent” the water body or a defined stratum thereof.

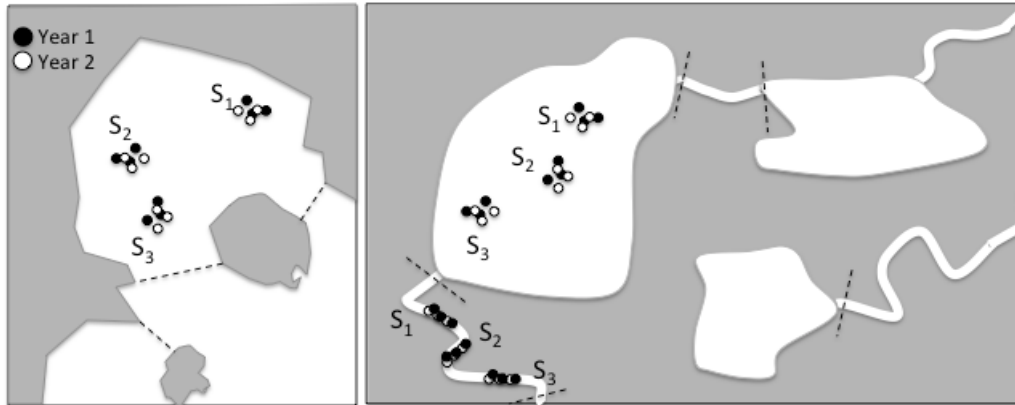


FIGURE 3.1
 Illustration of crossed monitoring designs in a coastal water body (left) and in a lake and stream (right). In the examples, $a = 2$ years, $b = 3$ stations, and $n = 3$ replicates. Reproduced from Lindegarth et al. (2013).

Each measurement made in such a programme may be expressed using a linear model in which the measured value, y , is the sum of the overall mean, μ , and deviations due to the other sources of variability.

$$y = \mu + YEAR + STATION + YEAR * STATION + PATCHINESS$$

The variability of the overall mean in such a sampling design consists of several variance components, i.e., s_Y^2 (variability among years), s_S^2 (variability among sites), s_{Y*S}^2 (changes in spatial variability across years), and s_e^2 (variability among replicates), each associated with a different source of variability in the linear model. The variance of the estimated mean, \bar{y} , resulting from these components in a crossed design can be calculated as:

$$V[\bar{y}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_S^2}{b} + \frac{s_{Y*S}^2}{ab} + \frac{s_e^2}{abn}$$

where a is the number of years sampled, b is the number of stations sampled, and n is the number of replicates taken at each station and sampling time (Figure 3.1). This formula for error propagation indicates how individual uncertainty components are combined into a total variance estimate and, importantly, how the numbers of replicates, stations, and years affect the variance and uncertainty. Increasing the number of replicates reduces the uncertainty due to small-scale variability within stations and years, but does not affect the uncertainty caused by variability among years or stations. Monitoring at many stations reduces the uncertainty due to spatial and replicate variability, but does not affect the uncertainty due to temporal variability. Similarly, sampling several years reduces the uncertainty due to temporal and replicate variability, but does not affect the uncertainty due to spatial variability. Note also that if all years within an assessment period are sampled, i.e., $a = Y = 6$, all possible levels of the factor are sampled, implying that the

distribution over the six years (constituting the entire relevant population) is known (estimated) and therefore does not contribute any random variation.

One relevant extension to this structure is the inclusion of a seasonal factor, when monthly samples are taken. This is typically the case in phytoplankton monitoring, and the current phytoplankton biomass indicator in coastal waters is the mean over three summer months (June–August). This results in the following linear model:

$$y = \mu + YEAR + MONTH + STATION + YEAR * MONTH + YEAR * STATION + MONTH * STATION + YEAR * MONTH * STATION + PATCHINESS$$

This also means that additional variance components associated with months and several interactions may have to be accounted for, i.e., s_M^2 (variability among months), s_{Y*M}^2 (changes in the monthly pattern among years), s_{M*S}^2 (changes in the monthly pattern among stations), and s_{Y*M*S}^2 (variability among samples at the same station taken in the same year and month). The total variance of the estimated mean, \bar{y} , resulting from these components in a crossed design can be calculated as:

$$V[\bar{y}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_S^2}{b} + \frac{s_M^2 * (1 - \frac{c}{M})}{c} + \frac{s_{Y*M}^2 * (1 - \frac{ac}{YM})}{ac} + \frac{s_{Y*S}^2}{ab} + \frac{s_{M*S}^2}{cb} + \frac{s_{Y*M*S}^2}{abc} + \frac{s_e^2}{abcn}$$

where, in addition to the above nomenclature, c is the number of months sampled (of the M months used for the indicator; in the case of a summer mean $M = 3$), b is the number of stations sampled, and n is the number of replicates taken at each station and sampling time.

3.1.2 Nested design in a single water body

Another fundamental design that is potentially useful but not commonly used in aquatic environments in Sweden is one in which new sites (“stations”) are sampled each year (Figure 3.2). As in the previous example, the sites may have been selected completely at random in the water body or using criteria such as a narrow depth range, substrate, or distance from shore. The sites are selected to represent the water body or a defined stratum. Note also that the number of sites (b) and replicates at sites (n) may vary greatly among monitoring programmes.

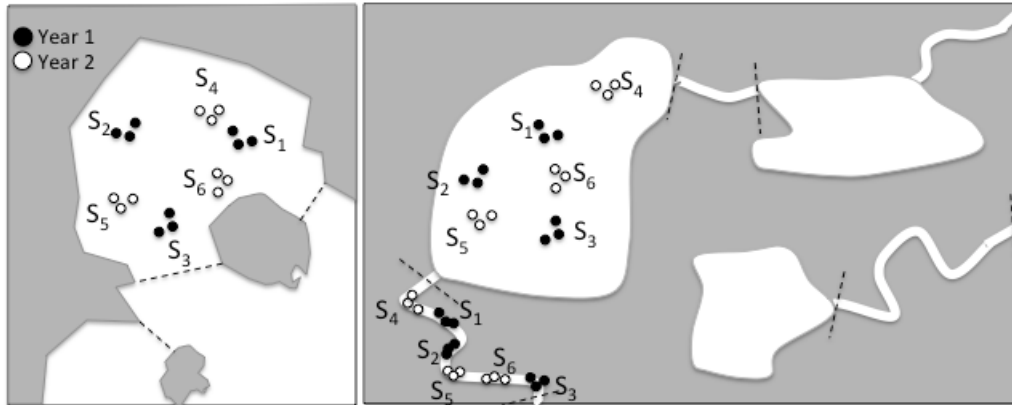


FIGURE 3.2

Illustration of nested monitoring designs in a coastal water body (left) and in a lake and stream (right). New sites are selected each year, so sites are nested within years. In the examples, $a = 2$ years, $b = 3$ sites, and $n = 3$ replicates. Reproduced from Lindegarth et al. (2013).

In this example, each measurement can be expressed using a linear model in which the measured value, y , is the sum of the overall mean, μ , and deviations due to the other sources of variability.

$$y = \mu + YEAR + SITES(YEAR) + PATCHINESS$$

The variability of the overall mean in such a sampling design consists of three variance components, i.e., s_Y^2 (variability among years), $s_{S(Y)}^2$ (variability among sites within years), and s_e^2 (variability among replicates), each associated with a different source of variability in the linear model. The variance of the estimated mean resulting from these components in a nested design can be calculated as:

$$V[\bar{y}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_{S(Y)}^2}{ab} + \frac{s_e^2}{abn}$$

This formula describes how the various uncertainty components contribute to the indicator variance as function of the numbers of replicates, sites, and years. Again, we can see that increasing the number of replicates reduces the uncertainty due to small-scale variability (patchiness), but does not affect the uncertainty caused by variability among years or sites. As in the previous example, sampling over a number of years will reduce the uncertainty due to interannual variability; if all years are sampled, the factor is considered completely fixed and the uncertainty component due to years becomes zero. One important difference from the crossed design is that the numbers of sites and years both contribute to reducing the spatial uncertainty, because sites are nested within years (i.e., new sites are measured every year). This may substantially reduce the indicator uncertainty if the spatial variability is much larger than the spatio-temporal variability, i.e., if there are consistent rather than transient differences among sites (i.e., $s_S^2 > s_{Y*S}^2$; Lindegarth et al.

(2013)). Another important difference is that the variance component, $s_{S(Y)}^2$, describes both the spatial variation across sites at any given time (s_Y^2) and the difference in this spatial variation across years (s_{Y*S}^2). Given the nested design, it is therefore impossible to partition $s_{S(Y)}^2$ further into the two other components.

Like the crossed design, monitoring designs for some quality elements may involve sampling on several occasions (months) across years. Each measurement in such cases can be expressed using a linear model in which the measured value, y , is the sum of the overall mean, μ , and deviations due to the other sources of variability.

$$y = \mu + YEAR + MONTH + YEAR * MONTH + SITES(YEAR * MONTH) + PATCHINESS$$

In these instances, each term in the model can be combined into a variance indicator (additional variance components denoted s_M^2 , s_{Y*M}^2 , and $s_{S(Y*M)}^2$):

$$V[\bar{y}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_M^2 * (1 - \frac{c}{M})}{c} + \frac{s_{Y*M}^2 * (1 - \frac{ac}{YM})}{ac} + \frac{s_{S(Y*M)}^2}{ab} + \frac{s_e^2}{abcn}$$

3.1.3 Crossed and nested designs in a water body type

The primary task of surveillance and operational monitoring is to provide a basis for status assessments of water bodies within a WFD cycle (see sections 3.1.1 and 3.1.2). Nevertheless, the uncertainty framework presented here can also be used to provide routines for estimating mean status and uncertainty in larger spatial units. For example, the aim of surveillance monitoring is to provide status assessments in both catchments and subcatchments. Alternatively, there may also be a need for routines for aggregating data from a number of water bodies into larger spatial units representing a certain water body type. This is in fact what is required for the Swedish MSFD assessment, which is intended to focus on water body types as the smallest assessment unit in coastal areas.

In such cases, the uncertainty framework can easily be extended to incorporate variability due to differences among water bodies. Each measurement in a programme may be expressed using a linear model, which includes the same components as those within a water body (see section 3.1.1), but with the addition of variability among water bodies (i.e., s_{WB}^2 and s_{Y*WB}^2). In a crossed design in which stations (and therefore water bodies) are revisited, the linear model and the total variance of the overall mean can be expressed as:

$$y_{Type} = \mu + YEAR + WATER BODY + STATION(WATER BODY) + YEAR * WATER BODY + YEAR * STATION(WATER BODY) + PATCHINESS$$

$$V[\bar{y}_{Type}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_{WB}^2}{b} + \frac{s_{S(WB)}^2}{c} + \frac{s_{Y*WB}^2}{ab} + \frac{s_{Y*S(WB)}^2}{abc} + \frac{s_e^2}{abcn}$$

If, on the other hand, the monitoring is designed as a nested sampling programme (see section 3.1.2) in which new sites are selected each year within the same water bodies (i.e., $S_{S(Y*WB)}^2$), the corresponding linear model and total variance is defined as:

$$y_{Type} = \mu + YEAR + WATER\ BODY + YEAR * WATER\ BODY \\ + SITES(YEAR * WATER\ BODY) + PATCHINESS$$

$$V[\bar{y}_{Type}] = \frac{s_Y^2 * (1 - \frac{a}{Y})}{a} + \frac{s_{WB}^2}{b} + \frac{s_{S(Y*WB)}^2}{abc} + \frac{s_e^2}{abcn}$$

3.2 Structure of current monitoring programmes

Biological monitoring in Sweden has been developed for many purposes. It is funded and organised by various national and regional authorities and has varying historical backgrounds. In addition to supporting ecological assessment in the WFD context, the aims of monitoring are mainly to provide information on progress towards meeting national environmental objectives, to ensure that Sweden complies with international conventions, to gather data for regional environmental impact assessments, and to evaluate the efficacy of management actions. Furthermore, the need for coordination with other EU directives, such as the Habitats Directive and the Marine Strategy Framework Directive, is becoming increasingly important.

Because the purpose and history of monitoring programmes is so diverse, their structure and size (and sometimes measured variables) often differ markedly. To evaluate the structure of programmes relevant to WFD status assessment, we collected information from the Swedish national database VattenInformationsSystem Sverige (VISS; <http://www.viss.lansstyrelsen.se>). This portal is run by the county administrative boards (Länsstyrelserna) and the Swedish River Basin District Authorities (Vattenmyndigheterna) and is a resource in which status assessments for all water bodies and quality elements can be found. VISS also contains geographical information on the network of monitoring stations potentially available for WFD status assessments and on the sampling frequency at all of these stations. In October 2012, information on sampling stations per water body and their sampling frequencies was extracted from VISS. This information was summarized and used to analyse fundamental structural properties and quantitative aspects of each BQE. Information on the number of samples per station and the sampling time, which is not given in VISS, is based on personal communications with WATERS collaborators and generally follows the national monitoring standards (see <https://www.havochvatten.se/kunskap-om-vara-vatten/datainsamling-och-miljoovervakning/handledning-for-miljoovervakning/undersokningstyper-och-miljoovervakningsmetoder.html>).

TABLE 3.1

Summary of monitoring data potentially available for WFD status assessment per BQE and water body. Note that the data from all sampled water bodies may not fulfil the requirements specified in the Swedish assessment criteria (NSF 2008:1). Numbers reflect the most common metric for BQEs for which several metrics are defined. See text and Figures 3.3–3.5 for more details on typical numbers of times, stations, and samples.

BQE	No. stations sampled	No. water bodies sampled [§]	% of all water bodies [§]	Typical no. times per WFD cycle	Typical no. stations per water body	Typical no. samples per station and time
Coastal waters						
Benthic invertebrates	155	77	12.8	6	1	1–5
Macrophytes	156	77	12.8	6	1	1
Phytoplankton	240	160	26.6	30	1	1
Lakes						
Benthic invertebrates	237	192	2.7	2	1	5
Macrophytes	49	48	0.7	1	1 (>8*)	1
Phytoplankton	459	459	6.3	6	1	5*
Fish	215	204	2.8	1	1 (8–64*)	1
Streams						
Benthic invertebrates	731	629	4.0	2	1	5
Benthic diatoms	413	376	2.4	6	1	5*
Fish	1056	645	4.1	6	1	1

* Data are pooled across stations or samples to calculate the Swedish WFD metric.

[§] During the preparation of this report, the River Basin District Authorities published a report on the monitoring of Swedish water bodies (Kronholm et al. 2013). The percentage of monitored water bodies reported there deviates slightly from the numbers presented here (probably due to later additions in VISS). Nevertheless, the focus here is on programme structure, which is unlikely to have changed in any substantial way.

3.2.1 Coastal waters

Swedish coastal waters are divided into 602 water bodies (Mårtensson et al. 2011). Benthic invertebrates and macrophytes are sampled in approximately 13% of these (Table 3.1). The number of stations per water body ranges between 1–12 for benthic invertebrates and 1–20 for macrophytes, but for both of these BQEs, 80% of the water bodies are represented by only one or two stations (Figure 3.3). Phytoplankton (chlorophyll a) is sampled in approximately 27% of the coastal water bodies. Even though approximately 5% of the water bodies have more than two stations, the number of stations per water body tends to be smaller for phytoplankton than for the other BQEs (Figure 3.3). Another aspect of the spatial sampling is the number of replicate samples taken at individual stations. Because traditional trend monitoring has often focussed on obtaining appropriate

representations of individual stations, some stations in programmes for benthic invertebrates yield several samples per station (4–5 on the Swedish west coast). Newer programmes focus more on obtaining samples representative of areas rather than stations. These stations typically yield one (in the national programme in Bothnian Bay and Baltic proper) or two (in the national/regional programme in Skagerrak) samples per station and year. Sampling of macrovegetation and plankton are predominantly done with one sample per station and year.

The temporal representativity of sampling differs strongly between the benthic BQEs and phytoplankton (Figure 3.3). The former are typically sampled once every year and thus six times during a WFD cycle. Nevertheless, a substantial proportion of the water bodies (30–40%) are sampled every second or every third year, resulting in two or three samples during a WFD cycle. Stations for phytoplankton, however, are sampled at least twice each year (i.e., twelve times each cycle), but more commonly 3–12 times per year (18–72 per cycle). This large difference in sampling frequency among benthos and plankton naturally reflects differences in the temporal dynamics of these variables, the fact that the monitoring of benthic flora and fauna is restricted to certain times of the year, and probably also differences in monitoring costs among different BQEs. Despite the seemingly frequent sampling for phytoplankton, it is important to note that only data from June to August are used to calculate current metrics.

One typical feature of all coastal BQEs is that sampling designs are almost exclusively crossed with respect to stations and sampling times, i.e., the same stations are revisited at each sampling time. This design is clearly a result of efforts to minimize uncertainty in estimates of temporal trends due to spatial variability.

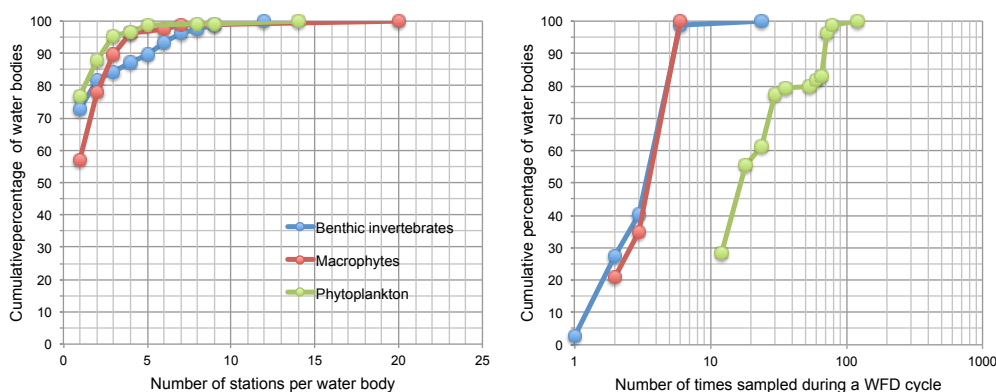


FIGURE 3.3

Cumulative distributions of the number of stations (left) per water body and sampling times and (right) per water body and six-year period in coastal areas. Number of stations defined according to the definitions in VISS (i.e., stations = “EU_CD övervakningsstation”).

These analyses indicate that the current monitoring of invertebrates in coastal water bodies typically involves sampling every year during a WFD cycle ($a = 6$). On each occasion, one station is sampled repeatedly ($b = 1$) and one to five core samples ($n = 1-5$) are taken (Table 3.1). Using the formulae described in section 3.1.1, this means that the total variance around the overall mean for an assessment period can be calculated as:

$$V[\bar{y}_{invertebrates}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6*1} + \frac{s_e^2}{6*1*1} = 0 + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6} + \frac{s_e^2}{6}, \text{ if } n = 1$$

or

$$V[\bar{y}_{invertebrates}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6*1} + \frac{s_e^2}{6*1*5} = 0 + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6} + \frac{s_e^2}{30}, \text{ if } n = 5.$$

Similarly, the monitoring of macrophytes in coastal water bodies typically involves sampling every year during the WFD cycle ($a = 6$). At each time, one station is sampled repeatedly ($b = 1$) and one transect is taken ($n = 1$). This means that the total variance around the overall mean for an assessment period can typically be calculated as:

$$V[\bar{y}_{macrophytes}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6*1} + \frac{s_e^2}{6*1*1} = 0 + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6} + \frac{s_e^2}{6}.$$

The monitoring of phytoplankton differs from that of the previous BQEs in that it typically involves sampling several times per year and in that the assessment criteria recommend that the yearly mean be based on the average of three summer measurements. Nevertheless, phytoplankton sampling typically involves sampling every year during the WFD cycle ($a = 6$). At each time, one station is sampled repeatedly ($b = 1$) and one sample is taken ($n = 1$). If all summer months are sampled (i.e., $c = 3$), the total variance around the overall mean for an assessment period can typically be calculated as:

$$V[\bar{y}_{phytoplankton}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_M^2 * (1 - \frac{3}{3})}{3} + \frac{s_{Y*M}^2 * (1 - \frac{6*3}{6*3})}{6*3} + \frac{s_{Y*S}^2}{6*1} + \frac{s_{M*S}^2}{3*1} + \frac{s_{Y*M*S}^2}{6*3*1} + \frac{s_e^2}{6*3*1*1} = 0 + \frac{s_S^2}{1} + 0 + 0 + \frac{s_{Y*S}^2}{6} + \frac{s_{M*S}^2}{3} + \frac{s_{Y*M*S}^2}{18} + \frac{s_e^2}{18}.$$

3.2.2 Lakes

The Swedish register of surface waters, SVAR, defines 7232 lake water bodies (Mårtensson et al. 2011). Approximately 6% of these are sampled for phytoplankton, 2–3% for fish and benthic invertebrates, and less than 1% for macrophytes (Table 3.1). The vast majority of these water bodies are sampled at one station at each sampling time. For phytoplankton and benthic invertebrates, one station per water body is the rule (approximately 20% of the water bodies are sampled for benthic invertebrates at two or more stations; Figure 3.4). In terms of small-scale replication within stations, for all BQEs except benthic invertebrates, sampling at one station results in one index value, even if several samples are sometimes taken. This is because the methods used for monitoring fish, macrophytes, and phytoplankton prescribe that samples be pooled. This means that many sources of uncertainty are combined into a single component (note, however, that these pooled

samples can be considered representative of the lake as several replicate sites are sampled). This is not the case for benthic invertebrates, and five samples and replicate values of the metric are obtained at each station and sampling time.

With regards to temporal sampling, monitoring is generally less extensive in lakes than in coastal areas (Figure 3.4). The most frequently sampled BQE is phytoplankton, which is typically sampled every year and thus six times per WFD cycle. Fewer than 25% of the water bodies are sampled more than once per year. Samples of benthic invertebrates are taken at intervals ranging from every sixth year to every year, with a median of every third year (i.e., twice per WFD cycle). Fish and macrophytes are both typically sampled once per WFD cycle, but here the frequency varies from every tenth year to every year for fish and from every sixth year to every third year for macrophytes.

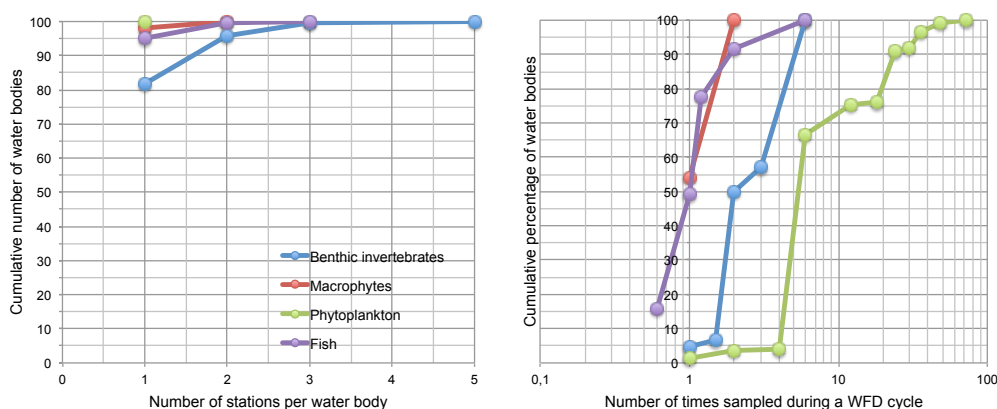


FIGURE 3.4

Cumulative distributions of the number of stations (left) per water body and sampling times and (right) per water body and six-year period in lakes. Number of stations defined according to the definitions in VISS (i.e., stations = “EU_CD övervakningsstation”).

Like coastal areas, lakes are generally sampled using crossed designs, i.e., stations are revisited year after year. Exceptions to this, however, are the monitoring programmes used to sample macrophytes and fish. In each lake, transects (macrophytes) or nets (fish) are placed at locations that are in principle revisited across years (though the exact placement and direction of the net may vary slightly). Nevertheless, these exceptions are of little importance because these metrics are calculated based on pooled samples for each lake and time.

The analyses indicate that the current monitoring of benthic invertebrates in lakes typically involves sampling twice during a WFD cycle ($a = 2$). At each time, one station is sampled repeatedly ($b = 1$) and five core samples ($n = 5$) are taken (Table 3.1). Therefore, the total variance around the overall mean for an assessment period can typically be calculated as:

$$\begin{aligned}
 V[\bar{y}_{invertebrates}] &= \frac{s_Y^2 * (1 - \frac{2}{6})}{2} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{2 * 1} + \frac{s_e^2}{2 * 5} \\
 &= \frac{s_Y^2 * (1 - \frac{2}{6})}{2} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{2} + \frac{s_e^2}{10}
 \end{aligned}$$

Monitoring of macrophytes in lakes typically involves sampling once per WFD cycle ($a = 1$). At each time, samples of the species composition are collected from different parts of the lake (“subjectively optimal”), ideally to produce a complete list of the macrophyte species present in the lake. A minimum of eight transects are used, but additional transects are sampled until the cumulative number of species levels out. Although the number of transects affects the uncertainty of the metric, the relationship between sample size (i.e., number of transects) and uncertainty cannot be assessed here. Nevertheless, the recommended sample size is likely based on knowledge of method-bound uncertainty; in NFS 2008:1, a rule of thumb is that when there is a deviation of 0.05 EQR units from a class boundary, a classification is considered “uncertain”. Although it is not explicitly stated, this is probably based on some kind of method-bound uncertainty (s_{Method}^2). However, given the recommendations in the Swedish assessment criteria and the uncertainty of estimates within a water body, the total uncertainty within a WFD cycle can be expressed as:

$$V[\bar{y}_{macrophytes}] = \frac{s_Y^2 * (1 - \frac{1}{6})}{1} + \frac{s_{Method}^2}{1}$$

Monitoring of phytoplankton in lakes is done by collecting one to five samples (depending on the lake size) once every year ($a = 6$) and is designed to estimate the status of the open water of the lake (i.e., affected as little as possible by benthic or littoral processes). Large lakes ($>1 \text{ km}^2$) are assumed to be well wind mixed, and in those, one sample from a central location in the lake is recommended. Small lakes are often wind protected by trees and are assumed to have a more patchy distribution of phytoplankton; accordingly, the recommendation is to pool five samples taken in the middle and each corner of a 200×200 -m square located in the central area of the lake. In practice, this means that samples can strictly only be considered representative of the status of a certain stratum of the lake, as the status in other parts of the lake may or may not differ substantially from that of the central part (i.e., the number of sites $b = 1$). Since samples are pooled into one sample, the sample size is effectively $n = 1$. With respect to the whole lake, the uncertainty can be expressed as:

$$V[\bar{y}_{phytoplankton}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6 * 1} + \frac{s_e^2}{6 * 1 * 1} = 0 + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6} + \frac{s_e^2}{6}$$

Regarding uncertainty, the monitoring of fish in lakes is structurally similar to that of macrophytes. Sampling programmes for fish in lakes are planned to give representative estimates of the indicator for the whole lake. Thus, 8–64 nets (depending on the lake size) are used, typically once per WFD cycle ($a = 1$), to assess the lake status. At each sampling time, the nets are placed at “semi-permanent” sites, i.e., the exact location and direction of

the nets are not completely identical among years. As samples are pooled to calculate the WFD metric, certain spatial components and those due to methodological errors are combined into one pooled uncertainty component:

$$V[\bar{y}_{fish}] = \frac{s_{\bar{y}}^2 * (1 - \frac{1}{6})}{1} + \frac{s_{Method}^2}{1}$$

3.2.3 Streams

SVAR defines 15,563 stream water bodies in Sweden (Mårtenson et al. 2011). Approximately 4% of these are sampled for benthic invertebrates, 2.4% for benthic diatoms, and 4% for fish (Table 3.1). According to the information available from VISS, more than 85% of these water bodies are sampled at one station each for benthic invertebrates and diatoms, while ~35% are sampled at more than one site for fish (Figure 3.5). The number of stations for fish sampling depends on the size of the drainage area and may vary among monitoring programmes. Five samples of benthic invertebrates and diatoms are taken at each station, and samples of invertebrates are kept separate whereas samples of diatoms are pooled. Electrofishing is done in streams by repeatedly sampling a defined stretch of water at least three times, to estimate the species-specific catchability needed to estimate the total number of fish in the sampled stretch.

Approximately 50% of the streams sampled for benthic diatoms and fish are sampled yearly, i.e., six times per WFD cycle (Figure 3.5). Benthic invertebrates are usually (~65%) sampled every third year. In line with programmes in coastal areas and lakes, monitoring is conducted regularly at fixed stations revisited on every sampling occasion.

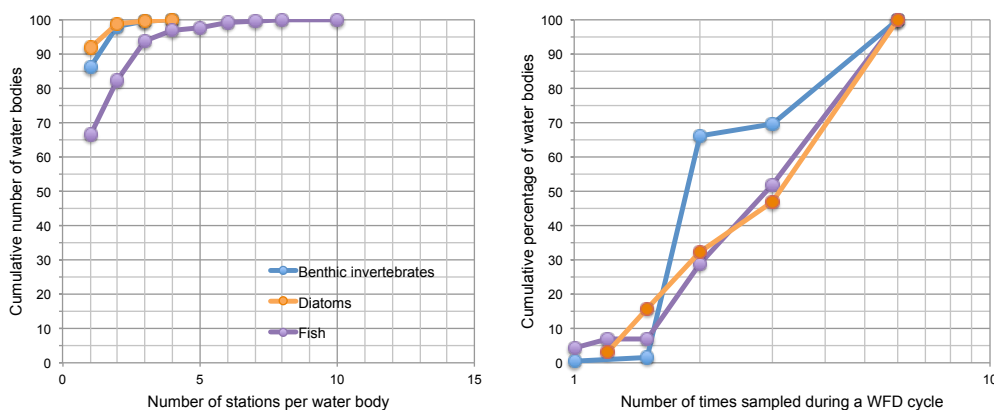


FIGURE 3.5

Cumulative distributions of the number of stations (left) per water body and sampling times and (right) per water body and six-year period in streams. Number of stations defined according to the definitions in VISS (i.e., stations = “EU_CD övervakningsstation”).

The typical structure of sampling within a water body for streams is exactly the same as for lakes (i.e., number of years, $a = 2$, number of stations per water, $b = 1$, and the number of samples per time and station, $n = 5$) and the total variance around the overall mean for an assessment period can be calculated as:

$$\begin{aligned} V[\bar{y}_{invertebrates}] &= \frac{s_Y^2 * (1 - \frac{2}{6})}{2} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{2 * 1} + \frac{s_e^2}{2 * 5} \\ &= \frac{s_Y^2 * (1 - \frac{2}{6})}{2} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{2} + \frac{s_e^2}{10} \end{aligned}$$

Benthic diatom monitoring in streams is similar to phytoplankton monitoring in lakes. Sampling is typically conducted by collecting five samples (stones) from a defined 10-m stretch of water once every year ($a = 6$). In 90% of the streams, one station is sampled (i.e., the number of sites $b = 1$); to reduce costs, the five samples collected at each time are pooled into one sample, meaning that the sample size is effectively $n = 1$. Regarding the stream as a whole, the uncertainty can be expressed as:

$$V[\bar{y}_{diatoms}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6 * 1} + \frac{s_e^2}{6 * 1 * 1} = 0 + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6} + \frac{s_e^2}{6}$$

The methods and structures of sampling programmes for fish in streams differ in many ways from those in lakes. In lakes, samples are taken using nets placed representatively throughout the lake, and the results of individual nets are pooled to calculate the WFD metric. Samples are typically taken once per WFD cycle. In streams, samples are taken by electrofishing at individual stations in a water body (usually at three stations in national monitoring programmes [$b = 3$], but often only at one station in regional programmes [$b = 1$]). Furthermore, samples are usually taken every year during the WFD cycle (i.e., $a = 6$). This means that the total uncertainty can typically be expressed as:

$$V[\bar{y}_{fish}] = \frac{s_Y^2 * (1 - \frac{6}{6})}{6} + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6 * 1} + \frac{s_e^2}{6 * 1 * 1} = 0 + \frac{s_S^2}{1} + \frac{s_{Y*S}^2}{6} + \frac{s_e^2}{6}$$

3.2.4 General features and conclusions

This review of the available monitoring programmes for quantifying WFD assessment indicators finds large structural diversity in spatial and temporal structure. As demonstrated, this has important consequences for the magnitude of uncertainty and the procedures for assessing uncertainty in coastal and inland water bodies. This review has aimed to analyse all BQEs separately, to provide insights into particular features and problems from the perspective of uncertainty assessment. It should be noted that the formulations focus on situations representative of each BQE, but that there are water bodies where the combinations of years, sites, and replicates differ slightly (for such instances, the general formulations in sections 3.1 need to be customised).

Despite the differences among BQEs and environments, a number of general features are evident:

- Few sites (or stations) are monitored per water body for all BQEs. The most common situation is that one site is monitored for each water body. This means that the variability among sites, and therefore the representativity, generally cannot be assessed. Unless this variability can be demonstrated to be generally small or negligible, this feature is likely to result in underestimation of the overall uncertainty. One possibility for evaluating the likely consequences of this deficiency is to quantify such spatial variability using existing data (i.e., water bodies where more than one site is monitored) or to target such sources of variability in specifically targeted sampling programmes, as discussed in section 4.
- Some BQEs usually employ sampling designs that do not allow for separation of the relevant sources of variability (e.g., due to lack of replicate samples, stations, and years within assessment periods). Because of this, it may be difficult to obtain sufficient information to optimise these programmes. Such problems may be addressed using the procedures discussed in section 4.
- The typical approach for most BQEs is to use designs in which sites are crossed with years, i.e., within a water body, the same sites being revisited at all sampling times. Because such designs can be used to “factor out” spatial from temporal variability, they are particularly appropriate for analysing temporal trends at individual stations. In terms of providing a spatially representative status assessment of, for example, a water body or water body type, such designs may be less effective than a nested design in which new sites are sampled at each time.

Finally, it is worth noting that these analyses are done with a primary focus on uncertainty in individual water bodies. The uncertainty of status assessments for water body types or catchments is probably less problematic, because individual sites in a number of water bodies can provide a representative sample of the water body type (see section 3.1.3). However, this requires that, following random or haphazard selection, sites be reasonably representative of their type from a defined part of the environment. The extent to which such an assumption is fulfilled can and should be evaluated using available environmental information.

4 Sampling designs to quantify uncertainty components

Monitoring programmes are generally designed so as to assess status and trends for a particular ecosystem, i.e., water body *sensu* the Water Framework Directive. Many different uncertainty components affect monitoring data (Lindegarth et al. 2013). Several of these uncertainty components cannot be estimated from regular monitoring data, whereas others cannot be identified independently but can only be estimated in combination with other sources of uncertainty. However, it is important to consider all relevant uncertainty components, and not just a subset of them, when assessing the uncertainty of a BQE indicator. Those uncertainty components that cannot be assessed using ordinary monitoring data can be estimated by designing controlled experiments with the specific aim of estimating these components.

One objective of the EU's WISER project (<http://www.wiser.eu/>) was to quantify various uncertainty components through field campaigns specifically designed for this purpose. The field campaigns were designed according to general guidelines that outlined a purely hierarchical design. Following this design, Dromph et al. (2013) analysed the variance in a number of phytoplankton pigments among water bodies, stations within water bodies, samples within stations, sub-samples within samples, and replicates, and found that for most pigments, variation among water bodies was the largest, followed by variation among stations. They did not, however, analyse for the temporal sources of uncertainty. Carvalho et al. (2013) analysed six phytoplankton metrics using a large pan-European lake dataset and similarly found that variation among water bodies was much larger than the spatial variation within lakes. However, they too did not investigate the temporal sources of uncertainty. Large spatial variability among lakes, sampling stations within lakes, and transects within stations was also found for macrophyte indicators using the WISER hierarchical design (Dudley et al. 2013). Finally, Balsby et al. (2013) used a large monitoring dataset of eelgrass depth limits from Denmark to quantify spatial, temporal, and methodological sources of variation. Overall, the variation among transects was the largest, but variations in diver-assessing depth limit among years and among replicates were of almost similar magnitude. Balsby et al. (2013) further demonstrated that the magnitude of these random components was depth dependent.

In this section, we will briefly discuss the principles of sampling designs in the context of the WFD, and demonstrate how various sampling designs can be used to quantify uncertainty components associated with spatial, temporal, and methodological variability.

4.1 Principles of sampling designs

There is no unique sampling design that should always be applied; in fact, uncertainty components can be estimated from various designs. The most important thing, however, is that the data be analysed using a linear statistical model that corresponds to the sampling design. A hierarchical design must be analysed using a hierarchical model; similarly, a crossed design and designs that combine hierarchical and crossed factors must be analysed using their corresponding models.

To illustrate the consequences of employing an appropriate versus an inappropriate statistical model to analyse data from a crossed versus a nested design, two datasets (crossed versus nested) were simulated with three random factors, A (variance = 1), B (variance = 4), and C (variance = 0.25). For the hierarchical dataset, factor B was nested within factor A, and factor C was the replicate variance for combinations of A and B (crossed design) or for combinations of B within A (nested design). The analysis was repeated 1000 times and the average variance components were calculated (Table 4.1). As expected, data analysed using correct models yielded estimates of the variance components close to the expected values, whereas choosing the wrong model yielded erroneous estimates.

TABLE 4.1

Average of estimated variance components from 1000 simulations of a crossed versus a nested design analysed with a crossed versus a nested model.

Simulated	Crossed design		Nested design	
	Crossed design	Nested design	Crossed design	Nested design
V(A)	0.992	0.609	1.158	0.994
V(B)	4.017	4.001	0.166	4.010
V(C)	0.250	0.250	3.939	0.250

In general, the choice of optimal sampling design is determined by the objective of the study (i.e., the specific statistical hypothesis) and economic or logistic constraints. Crossed designs, in which the same stations are visited repeatedly, have traditionally been preferred in environmental monitoring, because the objective has been to describe changes over time without considering whether the estimated trend was necessarily representative of the actual observation unit (i.e., water body *sensu* WFD). Hence, crossed designs are sub-optimal for comparing different spatial units. However, if the objective is the spatial comparison of water bodies, a hierarchical design would typically be chosen to reduce the spatial uncertainty, although such designs and analyses frequently ignore the temporal components. The monitoring objectives of the WFD are two-fold: 1) to provide precise status assessments of a given water body (suggesting a nested design to reduce the effect of spatial variability) and 2) to document improvement over time if a water body fails to meet environmental objectives (suggesting a crossed design to obtain a better spatial coupling of observations across assessment periods). An optimal design will likely include

a combination of maintaining permanent stations subject to repeated sampling and introducing new stations to improve indicator precision by reducing the effect of spatial variability.

4.2 Quantifying uncertainties not estimated by monitoring programmes

The uncertainty framework presented by Lindegarth et al. (2013) outlined a broad range of spatial, temporal, spatio-temporal, and methodological uncertainty components. It is unrealistic to design a sampling programme to be able to estimate all these uncertainty components, since this would require an enormous sampling effort if all components were to be sampled at a reasonable number of levels (i.e., replicates of the given factor). For example, sampling combinations of multiple stations with spatial replicates over multiple years and at different times of the year, using different people and sampling methods, can be challenging if the sole aim is to quantify magnitudes of uncertainty. Therefore, targeted monitoring programmes should be devised to estimate the specific uncertainty components for which the magnitude of random variation is not known but is supposedly important. For example, data on benthic flora and fauna do not fluctuate at short time scales, so random variation associated with diurnal fluctuations or short-term irregular fluctuations can be neglected.

One potentially important tool for quantifying other important sources of variation, i.e., spatial and temporal sources of variation, might be to add samples to existing programmes in a strategic way that allows estimation of these components without creating a fully balanced and complete design at all spatial and temporal scales. Such reduced designs can, under some circumstances, be cost-effective alternatives, provided that proper care is devoted to obtaining independent and representative estimates.

4.2.1 Estimation by reduced designs

To illustrate a potentially useful reduced design that might be used to estimate different sources of spatial variability, consider a situation in which the task is to assess the status of a BQE in a water body type (or a single water body) at one sampling time (the example can of course be extended to six-year periods, but for simplicity we concentrate on one year). The uncertainty of a water body type is affected by three sources of spatial variability: variability among water bodies (s_{WB}^2), variability among sites ($s_{S(WB)}^2$), and small-scale patchiness (s_e^2).

$$V[\bar{y}_{Type}] = \frac{s_{WB}^2}{a} + \frac{s_{S(WB)}^2}{b} + \frac{s_e^2}{abn}$$

To assess the uncertainty of water body types and water bodies, all of these components need to be estimated using appropriate statistical methods (e.g., ANOVA or maximum likelihood methods). As demonstrated earlier (e.g., section 3.2), the spatial structure of existing monitoring programmes varies among BQEs and environments, but quite often

the typical structure does not allow the estimation of all of these uncertainty components. For example, one crucial and common feature of several monitoring programmes is that they involve sampling at only one site per water body. This means that the potentially important variability among sites cannot be estimated and that the uncertainty cannot be appropriately assessed for the water body (Figure 4.1). The exact details vary among BQEs: in some instances one sample (sometimes pooled) is taken at each site (here called type A), whereas in others replicate samples are taken at each site (type B).

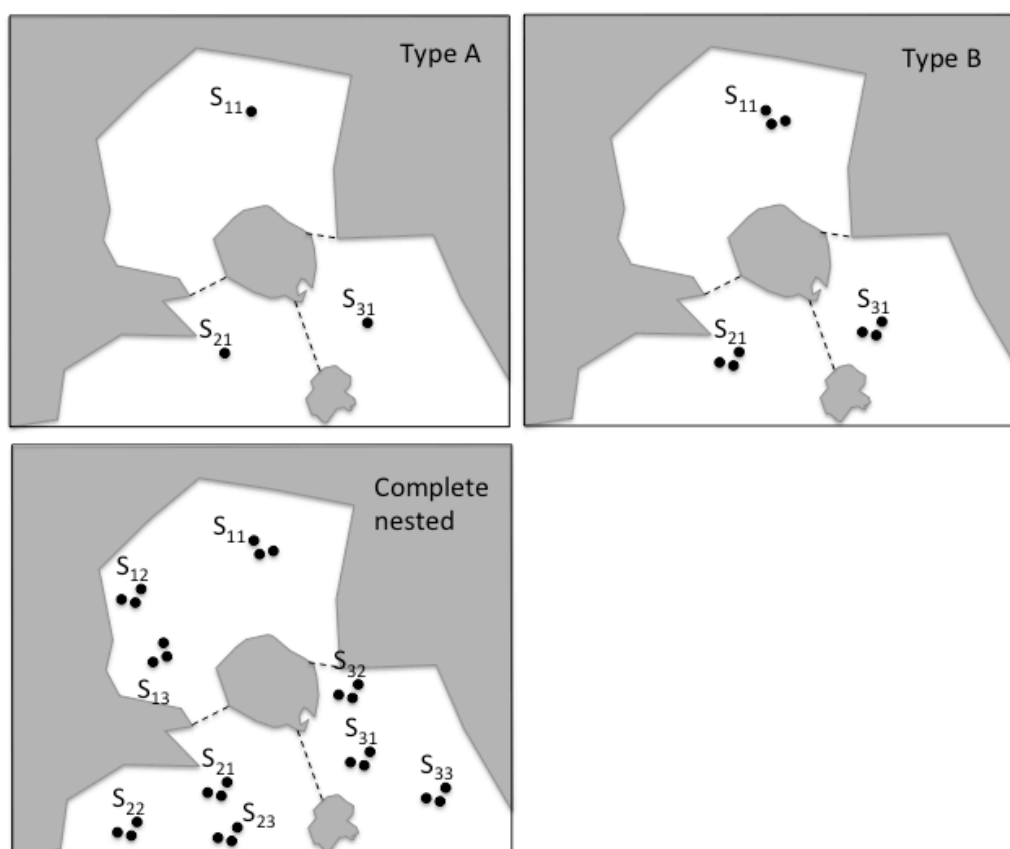


FIGURE 4.1

Schematic of the spatial arrangement of three monitoring designs with varying numbers of sites and replicates in three water bodies ($S_{\text{water body number, site number}}$). Types A and B are frequently seen in the current monitoring program, but such designs cannot be used to estimate all relevant sources of variation. The complete (“nested”) design might be used for such analysis.

The problem with these designs is that they do not allow assessment of how representative the samples are of their originating water bodies – in other words, we cannot estimate important uncertainty components. Type A is, of course, the least expensive design, but it only allows us to estimate the variability among water bodies (S_{WB}^2

or, strictly speaking, the combined uncertainty of $s_{WB}^2 + s_{S(WB)}^2 + s_e^2$), while type B allows calculation of s_{WB}^2 and s_e^2 ; see Table 4.2).

TABLE 4.2

Number of degrees of freedom (*df*) available to estimate uncertainty components (ne = not estimable) using four different sampling designs. See text and Figures 4.1 and 4.2 for explanation of the designs.

Uncertainty component	Type A	Type B	Complete nested design	Reduced staggered design	“Even” staggered design
Water body, = WB	2	2	2	2	2
Site, = S(WB)	ne	ne	6	6	12
Residual	ne	6	18	6	12
Total samples	3	9	27	15	27

To estimate the different sources of variation associated with water bodies, sites, and patchiness, it is necessary to have replicates at a number of spatial scales. A complete solution to this problem is to design monitoring that involves balanced replication at all spatial scales (i.e., *b* sites in each of *a* water bodies and *n* replicates in each of the *a*b* sites). The example shown in Figure 4.1 illustrates a case in which $a = b = n = 3$, but clearly it can be extended to any combination of numbers. This design is called a nested (hierarchical) design (e.g., Underwood 1997). Such a design can, in principle, be used to estimate all sources of spatial uncertainty, but two peculiarities of this design are that (1) it is expensive and (2) the number of degrees of freedom (*df*) is unevenly allocated among levels (Table 4.2). This means that some components are estimated relatively precisely while others are estimated less precisely. In this case, s_e^2 is estimated with $df = 18$ and $s_{S(WB)}^2$ with $df = 6$.

One type of design that might address the problems of excessive costs and the uneven allocation of *df* to the different components is a variant of a nested design called a “staggered” design (e.g., Khuri 2000, Ojima 2000). Staggered designs are unbalanced in that they have unequal numbers of samples within a hierarchical level (Figure 4.2). Such designs allow estimation of all components in a potentially more cost-effective way than do completely nested designs. For example, the resources needed to sample all the necessary scales may not be available unless the total number of samples is reduced (Table 4.2, Figure 4.2). Alternatively, the number of *df* available for estimating a component, in this case $s_{S(WB)}^2$, may not be large enough for sufficient precision. In such cases, samples may be allocated so that the number of *df* is similar for all relevant components (Table 4.2, Figure 4.2). In conclusion, it is easy to see how these designs may be used with data

originating from regular monitoring, strategically complemented with additional samples, to estimate important components of variability that ultimately can be used to assess uncertainty in water body types or water bodies.

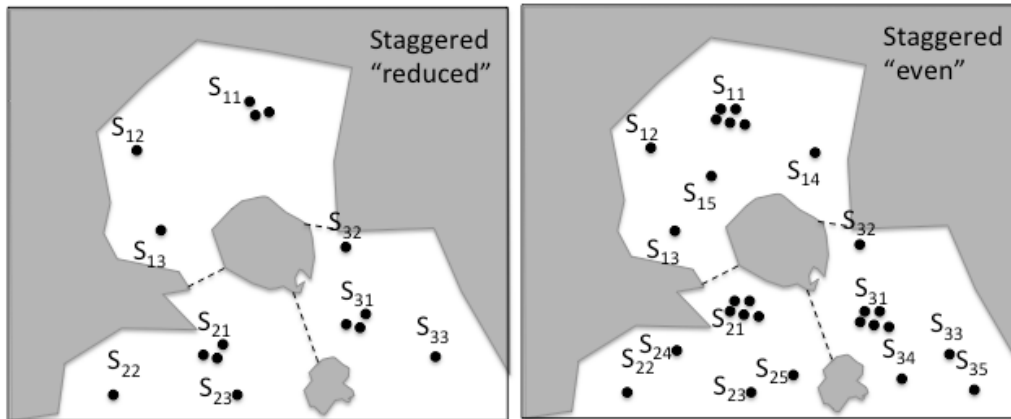


FIGURE 4.2

Schematic of the spatial arrangement of two examples of staggered designs that could be used to estimate most of the important spatio-temporal uncertainty components. The reduced design is less expensive in terms of number of samples than is the fully replicated design shown in Figure 4.1. The “even” design has the same number of samples as does the fully replicated design in Figure 4.1, but the degrees of freedom are more evenly distributed between the two lower spatial levels (see Table 4.2).

4.2.2 Estimation using information from other areas or times

Another way to alleviate potential problems of quantifying uncertainty components due to lack of replication at some level is to use information from other areas or times. We can exemplify the need to obtain additional information on sources of random variation using the BQI data from Lindegarth et al. (2013). The Skagerrak data were spatially replicated at the station and water body (or water body type) levels, so the random spatial variation could be estimated at both small and large scales (Table 4.3). There was no spatial replication at the station level in the Gulf of Bothnia, so the small-scale spatial variability could not be estimated separately from that among stations. This implies that the estimated variance component $V[\text{GRADIENT}]$ describes the combined small- and large-scale spatial variation. This combined random variation can be partitioned into the two uncertainty components by 1) collecting replicated samples at a number of stations and estimating the variation among replicates or 2) assuming that small-scale spatial variation in the Gulf of Bothnia is similar in magnitude to that in the Skagerrak. In the latter case, if we assume that $V[\text{REPLICATE}] \approx 0.5$, then this estimate can be subtracted from the estimate of the combined spatial variance to produce a variance estimate describing the large-scale variation only, i.e., for the area NAT1 this implies that $V[\text{GRADIENT}] = 4.080$ and $V[\text{REPLICATE}] = 0.5$.

TABLE 4.3

Spatial sampling efforts and estimated spatial variance components for BQI of benthic fauna in eight areas in the Skagerrak (2002–2004) and in the Gulf of Bothnia (1996–1998). Reproduced from Lindegarth et al. (2013).

Region	Area	No. stations	No. replicates	V[GRADIENT]	V[REPLICATE]
Skagerrak	Area 1	8	2	4.445	0.541
	Area 2	8	2	1.064	0.459
	Area 3	8	2	2.549	0.428
Gulf of Bothnia	NAT 1	22	1	4.580	-
	REG 2	20	1	3.881	-
	REG 4	20	1	6.244	-
	REG 6	20	1	8.242	-
	REG 8	20	1	3.126	-

However, the assumption that small-scale spatial variability in the Gulf of Bothnia is similar in magnitude to that on the west coast might not be supported, and a dedicated sampling effort could be initiated to investigate whether this indeed is the case or, alternatively, to provide a better estimate.

4.2.3 Data requirements for the estimation of uncertainty components

It is clear that the estimation of uncertainty components (i.e., variance components) is crucial for assessing the uncertainty of a status estimate or classification and for designing and optimising monitoring programmes. It is also important to realize that the estimates of these components are themselves subject to uncertainty. A general rule of thumb is that the greater the number of samples or, more correctly, the higher the *df*, available for estimating a certain component, the more precise the estimate is.

One way to assess the expected precision of an estimated component is to simulate the distribution and confidence interval of the component under different sampling schemes. To illustrate this approach, we assessed the potential data requirements for obtaining a reasonably precise estimate of the replicate variance from the example above (Table 4.3) in a simulation study assuming $V[\text{REPLICATE}] = 0.5$, using either eight stations as in the Skagerrak or 20 stations as in the Gulf of Bothnia, and varying the number of replicates taken at each station (Figure 4.3). With two replicates and eight stations, the estimated replicate variance from 16 observations could essentially be anything from 0.15 to 1.07. Increasing the number of replicates to 20 per station (i.e., 160 observations in total) reduced the expected range of the estimated replicate variance to 0.40–0.62. Increasing the number of stations to 20, as in the monitoring effort in the Gulf of Bothnia, resulted in a narrower range of estimated replicate variance, i.e., 0.26–0.87, with two replicates per station (40 observations in total). However, increasing the number of replicates per station to 20 substantially reduced the expected range of the estimated replicate variance to 0.43–

0.58. This example clearly illustrates the need for a sufficient number of replicates to estimate the variance components with adequate precision.

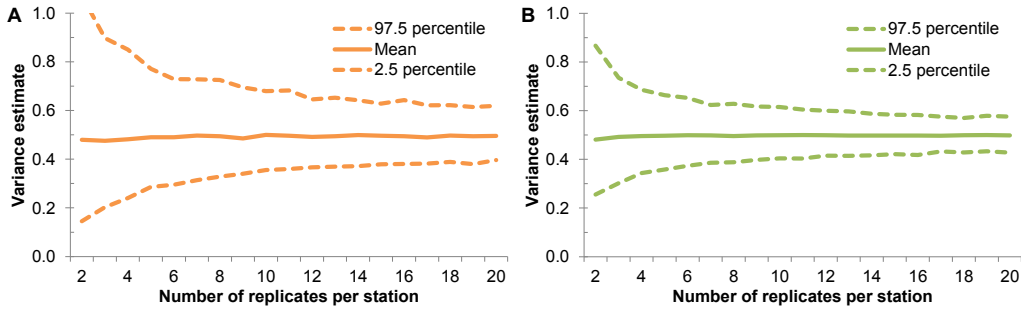


FIGURE 4.3 Distribution of estimated replicate variance (true value = 0.5) for increasing numbers of replicates per station obtained from 1000 simulations. A) based on eight stations similar to those in the Skagerrak region; B) based on 20 stations similar to those in the Gulf of Bothnia region (cf. Table 4.3).

The consequences of various sampling designs, particularly even more complex ones, can be investigated in a similar fashion by simulating the random distributions and estimating their variance components. The simulation approach is the most useful means of assessing data requirements and provides a useful test of a design before investing large resources in the actual sampling.

Another way of assessing the precision of an estimated variance is to use analytical formulae for the sampling distribution variances. It is well known that under assumptions of normality of the random variable X , the distribution of the estimated variance, S_X^2 , is a χ_{df}^2 distribution. Without going into details, the confidence interval of the true variance can be estimated from the following (e.g., Quinn and Keough 2002):

$$\frac{df * s_X^2}{\chi_{df, \alpha/2}^2} \leq \sigma_X^2 \leq \frac{df * s_X^2}{\chi_{df, 1-\alpha/2}^2}$$

where σ_X^2 is the true variance and $\chi_{df, \alpha/2}^2$ and $\chi_{df, 1-\alpha/2}^2$ are the critical values for the upper and lower percentiles, respectively, of the χ_{df}^2 distribution. Using this formulation, the confidence interval for the simulated distribution shown in Figure 4.3 can be reproduced (Figure 4.4). Furthermore, the relative size of the confidence interval (i.e., [upper limit–lower limit] s^{-2}) can be expressed as a function of df (Figure 4.4). This analysis demonstrates that the relative size of a 95% confidence interval of an estimated variance ≈ 1 at $df = 40$ and ≈ 0.4 when $df = 200$. However, the 95% confidence interval may perhaps be considered a strict criterion, and another way to formulate a target may be based on the average error of an estimated variance, i.e., its standard error, $SE_{S_X^2}$. It turns out that this quantity is a function of df and that it is calculated as:

$$SE_{s_x^2} = s_x^2 * \sqrt{2/df}$$

Finally, by dividing the standard error by the estimated variance, it is evident that the relative standard error can be calculated as $\sqrt{2/df}$. Using this expression, we can conclude that with $df \geq 22$ the average error is $\leq 30\%$ of the true variance, and to achieve an error $\sim 10\%$ of the variance, approximately 200 df are needed. Although there are no accepted targets or rules of thumb for what is sufficient precision for estimated uncertainties, it appears that a relative error of 30–20% is a reasonable target (i.e., $df = 30$ –50). In any case, these examples using simulation, analytical, and graphical tools illustrate useful approaches for determining the expected precision in the estimation of uncertainty components.

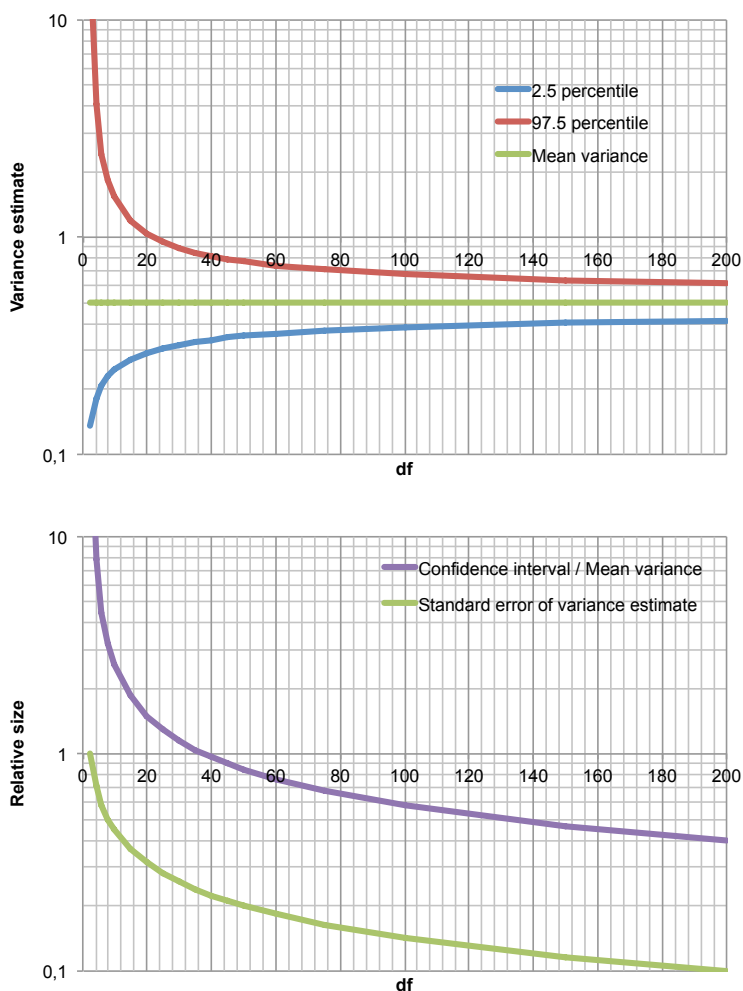


FIGURE 4.4 Uncertainty of estimated variance components using analytical expressions. Upper panel: Confidence interval of estimated variance ($\sigma^2 = 0.5$) with increasing df . Lower panel: Relative error expressed as confidence interval and standard error.

5 Conclusions

In this report, we have assessed conceptual issues related to monitoring BQEs for status assessment according to the WFD, and specific challenges associated with current monitoring using the uncertainty framework proposed by Lindegarth et al. (2013). The general conclusion is that the framework provides a useful tool for analysing uncertainties in a wide range of situations. The analyses have identified a number of general and specific properties of current monitoring that need to be addressed to ensure appropriate assessment and, ultimately, the reduction of uncertainty.

By analysing the typical structure of spatial and temporal replication within water bodies for each of the BQEs, we identified a large variety of monitoring approaches and, as a consequence, large differences in the combinations of relevant uncertainty components. A general concern for the assessment of individual water bodies is the frequent lack of replicate sites (or stations), causing a potential lack of spatial representativity. One important aspect, however, is that this lack of replication at the water body scale may not cause severe problems at the water body type or catchment scales, because assessment of status and uncertainty at these scales could be representative due to replication in a number of water bodies.

Recognising that there are historic, logistic, and, most importantly, economic constraints on monitoring and status assessment, we further illustrated how the uncertainty framework can be used in combination with existing data or the strategic addition of replicates at selected spatial scales in order to quantify critical components of uncertainty. For this purpose, we presented alternative designs, i.e., nested or staggered sampling designs, and illustrated methods for assessing the expected precision of estimates of variance components. For example, these methods indicate that the average deviation of an estimate from its true value (i.e., the standard error of the estimated variance) is 20–25% when the variance is estimated with 30–50 degrees of freedom. Although any rules of thumb for how precise a variance estimate needs to be are somewhat arbitrary, they do provide useful tools for WATERS' coming estimation of uncertainty components and for the compilation of an "uncertainty library", as suggested by Lindegarth et al. (2013).

6 Acknowledgements

We are grateful to members of the WATERS consortium for their constructive discussions and comments during the development of the uncertainty framework. Special thanks go to Kerstin Holmgren, Björn Bergquist, and Stina Drakare for their input regarding the text in general and, in particular, regarding portions of this report on inland fish and phytoplankton. The development of the report has also benefited from discussions with Ann-Karin Thorén and Patrik Börjesson, both at the Swedish Agency for Marine and Water Management, and with Juha Salonsaari at the River Basin District Authority.

7 References

- Balsby, T.J.S., Carstensen, J., Krause-Jensen, D. (2013) Sources of uncertainty in estimating eelgrass depth limits. *Hydrobiologia* 704: 311–323.
- Bennet, S., Roca, G., Romero, J., Alcoverro, T. (2011) Ecological status of seagrass ecosystems: An uncertainty analysis of the meadow classification based on the *Posidonia oceanica* multivariate index (POMI) *Marine Pollution Bulletin*, 62: 1616–1621.
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H., White, J.S. (2009) Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24:127–135.
- Carvalho, L., Poikane, S., Lyche Solheim, A., Phillips, G., Borics, G., Catalan, J. et al. (2013) Strength and uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes. *Hydrobiologia*, 704: 127–140.
- Clarke, R.T., Furse, M.T., Gunn, R.J.M., Winder, J.M., Wright, J.F. (2002) Sampling variation in macroinvertebrate data and implications for river quality indices. *Freshwater Biology*, 47:1735–1751.
- Clarke, R.T., Davy-Bowker, J., Sandin, L., Friberg, N., Johnson, R.K., Bis, B. (2006a) Estimates and comparisons of the effects of sampling variation using “national” macroinvertebrate sampling protocols on the precision of metrics used to assess ecological status. *Hydrobiologia*, 566:477–503.
- Clarke, R.T., Lorenz, A., Sandin, L., Schmidt-Kloiber, A., Strackbein, J., Kneebone, N.T., Haase, P. (2006b) Effects of sampling and sub-sampling variation using the STAR-AQEM sampling protocol on the precision of macroinvertebrate metrics. *Hydrobiologia*, 566:441–460.
- Clarke, R.T., Hering, D. (2006) Errors and uncertainty in bioassessment methods: Major results and conclusions from the STAR project and their application using STARBUGS. *Hydrobiologia*, 566:433–440.
- Clarke, R.T. (2013) Estimating confidence of European WFD ecological status class and WISER Bioassessment Uncertainty Guidance Software (WISERBUGS). *Hydrobiologia*, 704: 39–56.
- Cochran, William G. (1977) *Sampling techniques* (3rd ed.). Wiley, Hoboken, NJ, USA.

- Dromph, K., Agusti, S., Basset, A., Franco, J., Henriksen, P., Icely, J. et al. (2013) Sources of uncertainty in assessment of marine phytoplankton communities. *Hydrobiologia*, 704:253–264.
- Dudley, A., Dunbar, M., Penning, E., Kolada, A., Hellsten, S., Kanninen, A. (2012) Report on uncertainty in macrophyte metrics. WISER deliverable 3.2-2. <http://www.wiser.eu/download/D3.2-2.pdf>
- Dudley, A., Dunbar, M., Penning, E., Kolada, A., Hellsten, S., Oggioni, A. et al. (2013) Measurements of uncertainty in macrophyte metrics used to assess European lake water quality. *Hydrobiologia*, 704: 179–191.
- European Commission (2003) WFD CIS Guidance Document No 7. Guidance on the Intercalibration Monitoring under the Water Framework Directive. Produced by Working Group 2.7 – Monitoring. European Commission, Brussels. <http://circa.europa.eu/>
- European Commission (2000) Directive 2000/60/EC of the European Parliament and of the Council: Establishing a framework for Community action in the field of water policy. 23 October 2000. European Commission, Brussels.
- Khuri, A.I. (2000) Designs for variance components estimation: Past and present. *International Statistical Review* 68:311–322.
- Kronholm, M., Johansson, L., Denward, M., Öhrström, P., Lagergren, R., Nandorf, E. et al. (2013) Anpassning av övervakning till ramdirektivet för vatten: Vattenmyndigheternas förslag till strategi. Vattenmyndigheterna.
- Lindgarth, M., Carstensen, J., Johnson, R.K. (2013) Uncertainty of biological indicators for the WFD in Swedish water bodies: Current procedures and a proposed framework for the future. Deliverable 2.2-1, WATERS Report no. 2013:1. Havsmiljöinstitutet, Göteborg, Sweden.
- Mårtensson, J., Ranung, J., Asp, S.-S., Ceric, N., Gyllander, A., Johnsen, Å. et al. (2011) Arbete 2007–2010 med SVAR, Svenskt Vattenarkiv, en databas vid SMHI http://www.smhi.se/polopoly_fs/1.21181!SVAR-rapport_2010.pdf
- Mascaró, O., Bennett, S., Marbà, N., Nikolic, V., Romero, J., Duarte, C.M., et al. (2012) Uncertainty analysis along the ecological quality status of water bodies: The response of the *Posidonia oceanica* multivariate index (POMI) in three Mediterranean regions. *Marine Pollution Bulletin*, 64: 926–931.
- Ojima, Y. (2000): Generalized staggered nested designs for variance components estimation. *Journal of Applied Statistics*, 27:541–553.
- Quinn, G.P., Keough, M.J. (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge, UK.
- SEPA (2010) Status, potential and quality requirements for lakes, water courses, coastal and transitional waters: A handbook on how quality requirements in bodies of

surface water can be determined and monitored. Handbook 2007:4. Swedish Environmental Protection Agency, Stockholm.

Taylor, J.R. (1997) *An introduction to error analysis: The study of uncertainties in physical measurements*. University Science Books, Sausalito, CA, USA.

Underwood, A.J. (1997) *Experiments in ecology: Their logical design and interpretation using analysis of variance*. Cambridge University Press, Cambridge, UK.

Wikner, J., Kronholm, M., Sedin, A., Appelberg, M., Bignert, A., Leonardsson, K. et al. (2008) Strategi för kontrollerande övervakning av kust- vatten i Bottenvikens vattendistrikt. Report no. 2008:1. ISSN: 0348-029. Länsstyrelsen Västerbotten meddelande.

Monitoring biological indicators for the WFD in Swedish water bodies

Current designs and practical solutions for quantifying overall uncertainty and its components

In this report we use a comprehensive uncertainty framework for analysing and reviewing the monitoring requirements for BQEs, as defined in the EU Water Framework Directive, and the general spatial and temporal structure of existing monitoring in Swedish coastal and inland waters. The aims of the study are to illustrate the complexity of potentially important sources of uncertainty arising from the structure of monitoring, identify components of uncertainty that might need further attention and finally to suggest methods statistical and empirical methods for quantifying these components.

WATERS is coordinated by:



WATERS is funded by:



SWEDISH ENVIRONMENTAL
PROTECTION AGENCY