

1548

NINA Report

A New Sampling Protocol and Intercalibrated Index for Invertebrates in Running Water

Terje Bongard
Kristina Johansen
Tuva Bongard Munkeby



NINA Publications

NINA Report (NINA Rapport)

This is NINA's ordinary form of reporting completed research, monitoring or review work to clients. In addition, the series will include much of the institute's other reporting, for example from seminars and conferences, results of internal research and review work and literature studies, etc. NINA Report may also be issued in a second language where appropriate.

NINA Special Report (NINA Temahefte)

As the name suggests, special reports deal with special subjects. Special reports are produced as required and the series ranges widely: from systematic identification keys to information on important problem areas in society. NINA special reports are usually given a popular scientific form with more weight on illustrations than a NINA Report.

NINA Factsheet (NINA Fakta)

Factsheets have as their goal to make NINA's research results quickly and easily accessible to the general public. Fact sheets give a short presentation of some of our most important research themes.

Other publishing

In addition to reporting in NINA's own series, the institute's employees publish a large proportion of their scientific results in international journals, popular science books and magazines.

A New Sampling Protocol and Intercalibrated Index for Invertebrates in Running Water

Terje Bongard
Kristina Johansen*
Tuva Bongard Munkeby*

** Norwegian University of Science and Technology, 7491 Trondheim, Norway*

Bongard, T., Johansen, K., Munkeby, T.B. 2018. A New Sampling Protocol and Intercalibrated Index for Invertebrates in Running Water. NINA Report 1548. Norwegian Institute for Nature Research.

Trondheim October 2018

ISSN: 1504-3312

ISBN: 978-82-426-3286-9

COPYRIGHT

© Norwegian Institute for Nature Research

The publication may be freely cited where the source is acknowledged

AVAILABILITY

Open

PUBLICATION TYPE

Digital document (pdf)

QUALITY CONTROLLED BY

Carolyn Rosten

SIGNATURE OF RESPONSIBLE PERSON

Research director Ingebrigt Uglem (sign.)

CLIENT(S) REFERENCE(S)

Norwegian Environment Agency Contract no. 17011318.

COVER PICTURE

Extremely different river localities to be compared and assessed © NINA

KEY WORDS

- Water Frame Directive
- freshwater invertebrates
- ecological status index
- intercalibrated sampling protocol

CONTACT DETAILS

NINA head office

P.O.Box 5685 Torgarden
NO-7485 Trondheim
Norway
P: +47 73 80 14 00

NINA Oslo

Gaustadalléen 21
NO-0349 Oslo
Norway
P: +47 73 80 14 00

NINA Tromsø

P.O.Box 6606 Langnes
NO-9296 Tromsø
Norway
P: +47 77 75 04 00

NINA Lillehammer

Vormstuguvegen 40
NO-2624 Lillehammer
Norway
P: +47 73 80 14 00

NINA Bergen:

Thormøhlensgate 55
NO-5006 Bergen.
Norway
P: +47 73 80 14 00

www.nina.no

Abstract

Bongard, T., Johansen, K., Munkeby, T.B. 2018. A New Sampling Protocol and Intercalibrated Index for Invertebrates in Running Water. NINA Report 1548. Norwegian Institute for Nature Research.

The EU Water Framework Directive (WFD) needs a robust, reliable and practical method suitable for comparing the ecological status of invertebrate fauna in rivers. The ASPT index has several major shortcomings and needs to be revised. Obtaining statistically significant abundances of taxa requires time-consuming sorting, counting, and provides little extra information of value for assessment. We suggest collecting larger samples with focus on finding species, while avoiding counting whole samples. By subsampling taxa abundances in the process, we still obtained a useful approximation of numbers per minute sample. This approach is more robust to different sampling methods, personal effort and physical variables, provided sufficiently large samples are taken. Through this protocol, we propose an index called the Intercalibrated Benthic Invertebrate Biodiversity Index – IBIBI – based on each location's deviation from expected common species number. Deviance can be compared to WFD's ecological status scale, regardless of region or river type and is therefore intrinsically intercalibrated. In this study, we describe how the sampling protocol offers results of high quality comparable over time and space. Large datasets from Central Norway are used to exemplify the applicability of the protocol. We discuss our approach in the light of the 12 quality variables published in Bonada et al. (2006) and conclude that our protocol seems to meet all of the demands.

Terje Bongard¹, Kristina Johansen² and Tuva B. Munkeby²

¹Norwegian Institute for Nature Research, P.O. Box 5685, 7485 Trondheim, Norway

²Norwegian University of Science and Technology, 7491 Trondheim, Norway

Sammendrag

Bongard, T., Johansen, K., Munkeby, T.B. 2018. A New Sampling Protocol and Intercalibrated Index for Invertebrates in Running Water. NINA Rapport 1548. Norsk institutt for naturforskning

Vanndirektivet bruker en standard og en indeks (ASPT) som har betydelige usikkerheter og mangler. Vi foreslår her en ny tilnærming ved å diskutere flere forhold:

- Antall arter som registreres i en prøve fra en lokalitet er avhengig av prøvestørrelsen. Jo større prøve, jo flere arter vil registreres.
- De vanligste artene vil registreres først.
- Taksonomisk nivå bør være art, familienivå er for grovt, de fleste familier inneholder arter som både er følsomme og robuste mot påvirkninger.
- For å kunne interkalibrere, det vil si sammenligne lokaliteter over store avstander, foreslår vi å sammenligne hver lokalitet med sin egen opprinnelige, forventede artssammensetning. Dette avviket kan for hver lokalitet vurderes i forhold til de fem klassene for økologisk status beskrevet i vanndirektivet.
- Vi beskriver en utplukkingsmetode som er mer tidsbesparende, som kan ta større prøver på kortere tid. Plukkemetoden innebærer at en kan måle på nøyaktige antall per art og gruppe, men en får sett gjennom mye mer materiale og finner dermed flere arter.
- Vi presenterer forslag til forventningslister for Trøndelag og Møre, og viser noen eksempler på hvordan dette vil kunne anvendes for døgn-, stein- og vårfluefaunaen. Alle taksa kan imidlertid inkluderes eller brukes hver for seg på samme måte.
- Artslister over forventningssamfunn kan skreddersys til alle oppløsningsnivåer: Regioner med alle elvetyper, elvetyper innen regioner, kontinentalitet, høyde over havet og andre variabler kan tas hensyn til ved utarbeidelse av forventningslister.

Vi diskuterer protokollen i lys av de 12 kvalitetskravene som er publisert av Bonada m.fl. (2006), og konkluderer med at tilnærmingen ser ut til å tilfredsstille alle.

Terje Bongard¹, Kristina Johansen² and Tuva B. Munkeby²

¹Norwegian Institute for Nature Research, P.O. Box 5685, 7485 Trondheim, Norway

²Norwegian University of Science and Technology, 7491 Trondheim, Norway

Contents

Abstract	3
Sammendrag	4
Contents	5
1 Introduction.....	6
2 Material and methods.....	8
3 Results.....	11
3.1 Sampling Protocol and Handling time	11
3.2 The IBIBI index	11
4 Discussion.....	16
4.1 ASPT shortcomings.....	19
4.2 Requirements for a robust and practical index.....	20
5 Conclusion	23
6 References	24

1 Introduction

Functional freshwater ecosystems are essential parts of human life in many ways (Arthington, Naiman, McClain, & Nilsson, 2010; MEA, 2005; Ormerod, Dobson, Hildrew, & Townsend, 2010). Ecosystem services provided by freshwater range from drinking water and power production to regeneration of basic and vital living conditions for human life on earth. Benthic invertebrates are an important biological quality element used in assessing freshwater ecological status within the EU Water Framework Directive (Directive, 2000, 2003, 2005). It is vital to have a tool that is applicable and reliable in this assessment. Several projects have addressed issues associated with the practical implementation of the WFD (AQEM, STAR, REBECCA), see reviews and results in Furse et al. (2006). For instance, the STAR project has addressed two main problems of comparison: a) The adoption of identical sampling techniques across Europe, and b) the harmonization of how the national assessment systems are classified (Buffagni, Erba, & Furse, 2007; Buffagni & Furse, 2006; Erba et al., 2009). The Geographical Intercalibration Groups (GIGs) aimed to intercalibrate national index data to compare regions and countries across Europe (Bennett et al., 2011; Erba et al., 2009; European, 2008; M. L. McGarrigle & Lucey, 2009; Murray-Bligh, Buffagni, Cazzola, Birk, & Vlek, 2006). The complex problems of intercalibration and methodology are not yet solved, and the reliability of the results are still unclear (Buffagni & Furse, 2006; Furse et al., 2006). During the last years, two Norwegian projects have been performed with the purpose of trying to understand sampling variation and calibrate methods (Petrin et al., 2016; Velle et al., 2018, in Norwegian). The two most important lessons to be learned from these projects are that sampling effort is extremely difficult to standardize, and that counting abundances and species determination are correlated with time, money and resources. Sampling is influenced by each sampler's personal behavior, strength, movements and choices of sampling spots. A river bed substrate is a stochastic mosaic of rocks, gravel, organic debris and aufwuchs that notoriously yields different outcomes from parallel standardized sample sizes. As for the second lesson, competing over prize is becoming more and more important in monitoring projects. Here, we suggest a way to meet these challenges, and at the same time obtain a more accurate description of ecological status in accordance with the WFD.

An appropriate index should be comparable both over time and space, and at the same time anchored to the local reference condition. Different macroinvertebrate taxa have various levels of perturbation tolerance. These qualities are used in many indices, including The Biological Monitoring Working Party (BMWP) score, and the derived average score per taxon (ASPT) (Armitage, 1983; Walley & Hawkes, 1996, 1997). These indices are based on selected family taxa appointed to have general ecological demands. This is contrary to ecological knowledge. Species is the unit that should be used, both for tolerance and biodiversity purposes (V. H. Resh & Rosenberg, 1993). Species that constitute BMWP/ASPT-families and ultimately define the index benchmark are unevenly distributed across Europe. In Norway and Scandinavia, species distribution and abundances are different from Great Britain or Italy. Comparisons and intercalibration within Europe are consequently difficult to interpret (Aroviita, Mykra, & Hamalainen, 2010; Hawkins, Norris, Hogue, & Feminella, 2000).

We propose an index based on direct comparisons of localities with their own reference conditions, based on ecosystem, region or river class. We describe a sampling protocol for obtaining data suitable for this approach. By increasing sample size and focus on species detection instead of time consuming counting of abundances, deviance from expected EPT species number can be observed and used as an index. The result is an intercalibrated, feasible method of practical biomonitoring across regions in accordance with the five-scale ecological status classification of the WFD.

Species distributions and abundances in general follow a three-way pattern: 1. Some species are always rare regardless of a wide or endemic distribution pattern. 2. Some species may be locally common, but are rare over large areas, or: 3. Some species are common over large areas and regions. A number of species are common within a region even regardless of river type. On this basis, we suggest an index based on the Ephemeroptera, Plecoptera and Trichoptera (EPT) species predicted to be commonly present. This list forms a benchmark which can be adjusted to any area, a specific region, or a specific river type within a region. The topography, altitude, soil, water flow or size of regions may of course vary considerably. Nevertheless, the result is an

Ecological Quality Ratio from each specific locality, presented as deviance from a predicted species benchmark list. This deviance can be adjusted to the five-scale WFD classification. No extra intercalibration will be necessary, in that each locality's deviance from pristine conditions within the specific locality's own region will provide ecological status. We call this index approach the Intercalibrated Benthic Invertebrate Biodiversity Index (IBIBI). The index rests on two pillars (T. Bongard, Diserud, Sandlund, & Aagaard, 2011):

1. A sufficiently large sample size.
2. Predictions of common species present in the region, based on empirical knowledge.

2 Material and methods

We have based our results on two sources: Data from decades of sampling from Central Norway are available in the national freshwater invertebrate database “Vannmiljø” (<http://vannmiljo.miljodirektoratet.no/>, only in Norwegian). In addition to this data, we sampled 12 localities with the protocol described below (K. Aagaard, Bækken, & Jonsson, 2002; K. Aagaard & Dolmen, 1996; K. Aagaard, Solem, Bongard, & Hanssen, 2004; T. Bongard, 2005, 2009; T. Bongard & Aagaard, 2006; T. Bongard, Koksvik, J.I., 1989; Nøst, 1986; Ugedal et al., 2014). Twenty ten-minute kick samples from 12 rivers were collected. Four were situated around Trondheim city, Central Norway, collected during 2016 (Frost, Huni, & Kershaw, 1971). Eight localities were from both Trøndelag and Møre. About 130 000 organisms were handled in the process. 58 EPT species were recorded, respectively 15 Ephemeroptera, 17 Plecoptera and 26 Trichoptera. Sorting time and species determination time are presented for four of the localities as an example of time and resource effort invested (Appendix 2). No red-listed species were found.

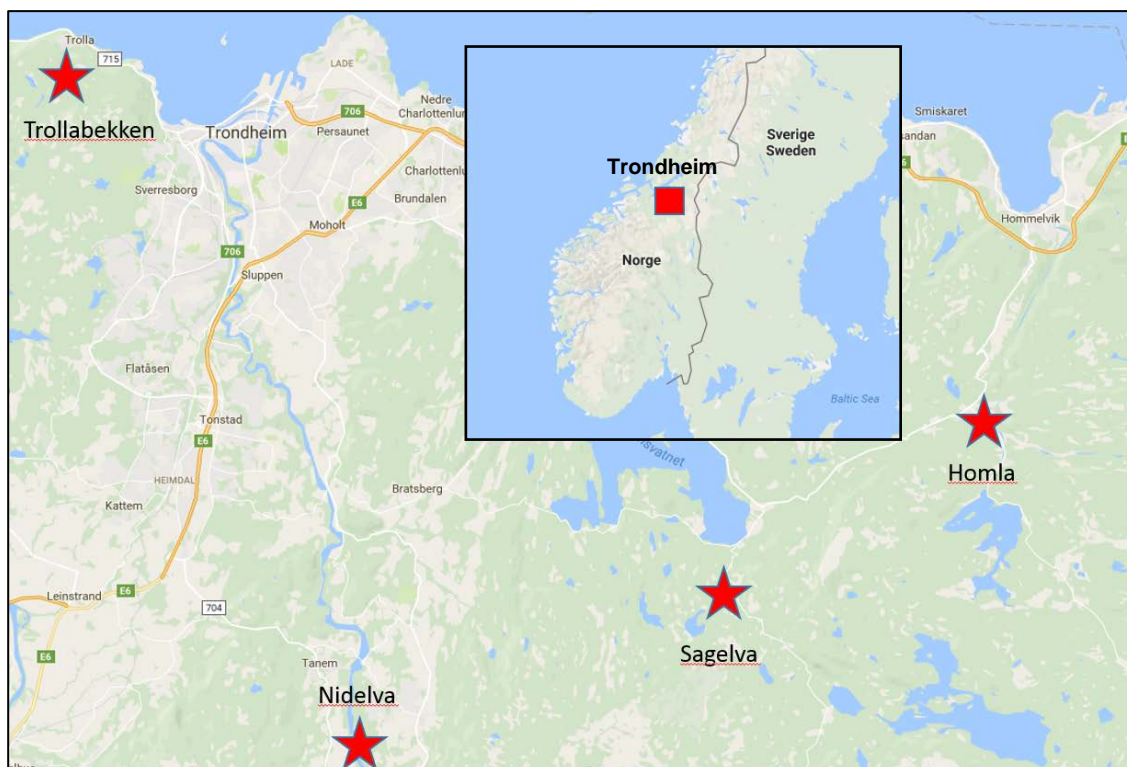


Figure 1. Sampling sites located around Trondheim, Central Norway. UTM references in Appendix 1.

The four localities around Trondheim (**Figure 1**) were kick sampled five times during 2016, from April to November (**Appendix 2**).

Net mesh size was 500 µm. To avoid clogging, two five-minute samples were combined into one. In rivers with more fine sediments or aufwuchs, a shorter sampling time might be necessary.

Samples were sorted with the equipment shown in **Figure 2**, which consists of a microscope with a 0.32 X objective and 10 X oculars, zoom 8-64 X, mounted on a boom arm stand. Such stands are produced in several versions, available for sale. A manageable amount of sample material was poured into a tray of size 43 x 25 cm, examined and sorted. The tray was repeatedly filled until the whole sample was examined. In the survey, live samples were sorted. Conserved samples may of course also be sorted with this equipment (Haase, Pauls, Sundermann, & Zenker, 2004).

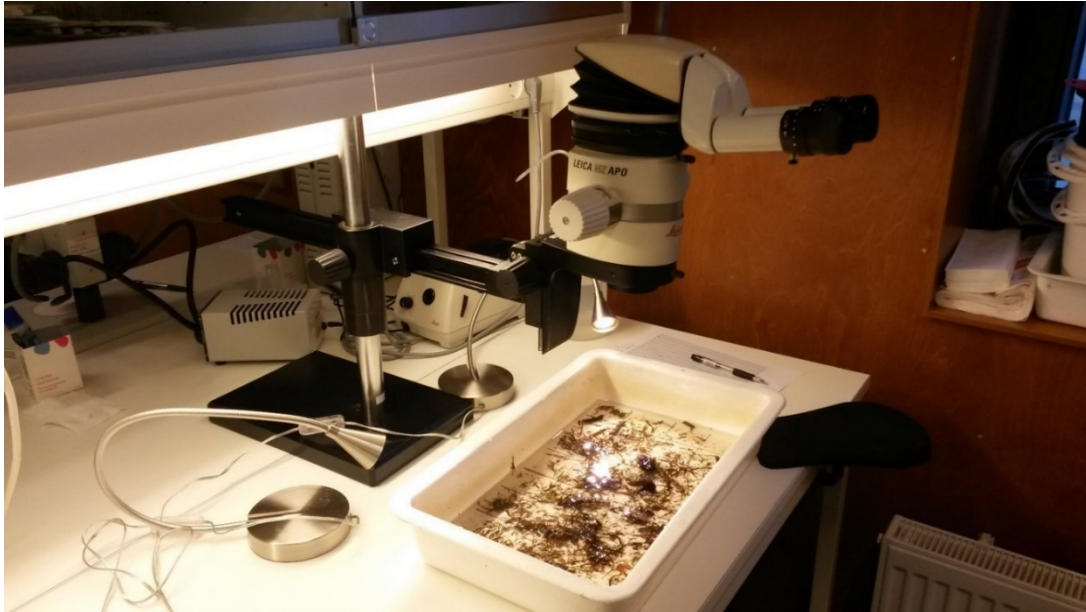


Figure 2. *Sorting equipment used. A microscope mounted on a sliding arm maintain stable focus all over the tray bottom. Note the marker grid on the tray bottom.*

The bottom of the tray was divided into 32 boxes of equal size with a permanent marker pen. Approximate abundances were obtained by counting taxa in a random selection of rectangles for each tray. An example of a data sheet produced by this method is shown in **Appendix 3**. This sample was one of the largest in the survey, containing about 8 liters of material. For this sample, 10 trays were examined.

Specimens of taxa not determined by this magnification were collected in vials with 70 % ethanol for later determination. Specimens from all taxa may of course be collected and preserved for different purposes if needed. Larger animals were picked with forceps, smaller were sucked up with a rubber bulb and squirted into a small piece of sieve net, allowing water runoff before transfer to a vial (**Figure 3**). This technique provides an efficient and time saving procedure.



Figure 3. *Picking equipment. Sucking organisms from the sample with a rubber bulb mounted on a 10 mm glass tube, and squirt them into a small net, allows for a much more effective sorting than forceps alone.*

Sorting efficiency with a microscope on a boom arm is a trade-off between area and magnification. The benefit of a large magnified area covering more of the sample will not provide sufficient magnification to determine all taxa. A zoom lens compensates this, and taxa that need higher resolution for determination are collected in a vial. Genera and species generally differ in their magnification need for proper determination. Difficult genera for species determination, like *Baetis*, *Rhyacophila* and *Hydropsyche*, consist of far more species in Central Europe than in Norway. Problems can be omitted by using such genera as benchmark taxa, see discussion for further elaboration on this.

After sorting, the microscope was placed on its original stand and the objective changed from 0.32 X to 1 X (Planapo) for higher magnification. Collected samples were transferred to a Petri dish for species determination. Mayflies, stoneflies, caddisflies, riffle beetles and alderflies were determined to species, other taxa to families.

3 Results

From this data, lists of expected EPT species in pristine conditions for Central Norway and North-West Norway were gathered (**Table 1**). Ecological status, as described in the WFD, are estimated as deviances from reference conditions. The outcome is two indices adapted to Central Norway and North-West Norway. Results from the 20 samples are shown in **Appendix 2**.

3.1 Sampling Protocol and Handling time

The equipment and procedure allow proper examination of large quantities of sample material within a relatively short time. The largest sample was about 8 litres, the smallest about 2 litres. Sorting time ranged from 0.5 up to two hours, including subsampling of abundances (**Appendix 2**, bottom row of table). Species determination of the five taxa included took 15 to 30 minutes per sample. No species hard to determine was found, which would have prolonged handling time. In all, total sorting time and species determination time for 200 minutes of kick sampling was less than 30 hours. Note that the protocol uses 500 µm net, with less clogging during sampling and easier sorting through less FPOM (fine particular organic matter). Collecting at several time points from spring to autumn allowed small specimens to grow until next sampling, making them easier to detect and determine.

3.2 The IBIBI index

Bongard et al (2011) introduced the idea of comparing samples to a predicted species composition from the same region. Deviations from expected species compositions can be scaled and adjusted to interpret the WFD ecological status definition in Annex V (Directive, 2003). Results will, through this process, be automatically intercalibrated, i.e. directly comparable across Europe. Lists of expected species for regions or water courses are open for adjustments over time, either because of new knowledge, or maybe to reflect any permanent changes in ecosystems. We present benchmark lists for two regions in **Table 1**. The lists are not based on specific knowledge of sensitivity or ecological demands for each species. The only criterion is that species are common and are expected to be present. Specific species knowledge concerning sensitivities may naturally be used and incorporated at any time.

Some species commonly have low abundances but are evenly distributed within microhabitat levels and therefore still have high probability for detection. This is the case for species like *D. nanseni*, *R. nubila*, *Potamophylax cingulatus*, *P. latipennis*, and *S. personatum*. Species such as *Athripsodes cinereus* may be equally abundant as *S. personatum*, but seem to be more patchily distributed within each microhabitat and are therefore not as detected readily as *S. personatum*. Such species are therefore not suited for our index.

Table 1. IBIBI Benchmarks: 27 species predicted to be present in medium sized, stony substrate pristine rivers in boreal, Central Norway, based on four sampling times from spring to autumn, 10 minutes of kicking per sample. Approximate expected abundances in a one-minute kick sample indicated for each species: * Usually few (Below 10); ** Usually common (10 to 50); *** Often dominant in numbers (more than 50). Total expected number of organisms in a one-minute kick sample is about 300-500.

Species	Approximate expected abundance per minute sample
Ephemeroptera:	
<i>Baetis muticus</i>	**
<i>B. rhodani</i>	***
<i>B. scambus</i>	**
<i>Heptagenia</i> spp.	**
<i>Ameletus inopinatus</i>	**
<i>Ephemerella aroni</i>	**
<i>E. mucronata</i>	*
Plecoptera:	
<i>Diura nanseni</i>	*
<i>Amphinemura</i> spp.	**
<i>Nemoura cinerea</i>	*
<i>Leuctra</i> spp.	**
<i>Taeniopteryx nebulosa</i>	*/**
<i>Isoperla obscura</i>	*/**
<i>Isoperla grammatica</i>	*/**
<i>Capnia</i> spp.	**
<i>Siphonoperla burmeisteri</i>	*/**
<i>Brachyptera risi</i>	*/**
<i>Protonemura meyeri</i>	*/**
Trichoptera:	
<i>Rhyacophila nubila</i>	*
<i>Hydroptila</i> spp.	*
<i>Polycentropus flavomaculatus</i>	**
<i>Apatania</i> spp.	*
<i>Potamophylax</i> spp.	*
<i>Arctopsyche ladogensis</i>	*
<i>Sericostoma personatum</i>	*
<i>Lepidostoma hirtum</i>	*
<i>Hydropsyche</i> spp.	*/**

Table 1 (cont.) 21 species expected to be present in medium sized, stony substrate pristine rivers in North-West Norway (Møre and Romsdal county), based on four sampling times from spring to autumn, 10 minutes of kicking per sample. Approximate expected abundances in a one-minute kick sample indicated for each species: * Usually few (Below 10); ** Usually common (10 to 50); *** Often dominant in numbers (more than 50). Total expected number of organisms in a one-minute kick sample is about 300-500.

Species	Approximate expected abundance per minute sample
Ephemeroptera:	
<i>Baetis muticus</i>	**
<i>B. rhodani</i>	***
<i>B. scambus</i>	**
<i>Ephemerella aroni</i>	**
<i>Leptophlebia</i> spp.	*
Plecoptera:	
<i>Diura nanseni</i>	*
<i>Amphinemura borealis</i>	**
<i>Leuctra</i> spp.	**
<i>Taeniopteryx nebulosa</i>	*
<i>Isoperla</i> spp.	*
<i>Capnia</i> spp.	**
<i>Siphonoperla burmeisteri</i>	*
<i>Brachyptera risi</i>	*/**
<i>Protonemura meyeri</i>	*
Trichoptera:	
<i>Rhyacophila nubila</i>	*
<i>Polycentropus flavomaculatus</i>	*
<i>Plectrocnemia conspersa</i>	*
<i>Potamophylax</i> spp.	*
<i>Oxyethira</i> spp.	*
<i>Sericostoma personatum</i>	*
<i>Lepidostoma hirtum</i>	*

From this list, we propose boundaries for the five ecological status classes (**Table 2**).

Table 2. Classification of ecological status for the investigated ecoregion in Central Norway, based on number of predicted EPT species present in pristine conditions (see Table 1).

Ecological Status	Number of EPT species
High	≥ 23
Good	18-22
Moderate	13-17
Poor	8-12
Bad	≤ 7

The four investigated localities in Central Norway and their deviances from this benchmark are presented in **Table 3**.

Table 3. Classification of ecological status for the investigated localities in rivers in Central Norway, based on deviations from number of predicted EPT species present in pristine conditions (Table 2).

Locality	Ecological Status
Homla	Good
Nidelva	Good
Sagelva	Good
Trollabekken	Moderate

Species diversity in Trollabekken was somewhat lower than expected (**Appendix 2**).

A classification of the two investigated localities in North-West Norway can be done in the same way. Proposed boundaries for the region are presented in **Table 4**, and results for the two rivers investigated is shown in **Table 5**:

Table 4. Classification of ecological status for North-West Norway, based on number of predicted EPT species present in pristine conditions (see Table 1).

Ecological Status	Number of EPT species
High	≥ 18
Good	15-17
Moderate	12-14
Poor	8-11
Bad	≤ 7

Table 5. Classification of ecological status for the investigated localities in North-West Norway, based on deviations from number of predicted EPT species present in pristine conditions (Table 4).

Locality	Ecological Status
Ålvunda	Good (17)
Søya	Moderate (13)

These results are in accordance with expectations. Ålvunda runs undisturbed from a National park, and only lower parts are regulated. Søya is canalized in most of its length and therefore has a high degree of anthropogenic impact.

4 Discussion

The two most commonly used methods for collecting invertebrates in rivers are the Surber and kick nets. There are relatively few published studies addressing the questions of sample size, number of parallels, subsampling techniques and sorting procedures in freshwater (Burdett, Fencil, & Turner, 2015; , see Feeley, Woods, Baars, & Kelly-Quinn, 2012 for a more thorough reference list; Haase, Lohse, et al., 2004; Haase et al., 2006; Haase, Pauls, et al., 2004; Haase, Pauls, Engelhardt, & Sundermann, 2008; Haase, Pauls, Schindehutte, & Sundermann, 2010; Stark, 1993; Storey, 1991; Vlek, Sporka, & Krno, 2006).

A river substrate is a micro-mosaic of stones, algae and micro stream velocities that leads to complicated sampling procedures, altered sampling performances and results, and consequently produces uncertain data. Currently, Norway uses a complicated protocol that describes how to work 20 seconds per meter for three times three meters, yielding a total sample of three minutes. It is practically difficult to perform such sampling. In practice, approximations burden all forms of sampling methods, standards and statistics, which depend on variables such as sampling time, substrate type, clogging, water velocity, depth, and individual differences in sampling performances (Carter & Resh, 2001; V. H. Resh & Rosenberg, 1993). The alluring and captivating concept of sample size standardization needs to be addressed and discussed more thoroughly. A common example of the effect of current sampling procedures on sample content is sampling from a stony river with a kick net. The depth, current and amount of vegetative matter collected by the net constantly change in the process. A stone is stuck, and ten seconds of struggling with it might be successful, and if so yields a load of matter. If the stone remains stuck, the time passes without any result. Individual judgment of how much time to use on such stones, or merely go around, considerably alters sampling results, both within and between localities.

Likewise, a Surber sampler pushed down on a stony substrate reveals how vulnerable this method is. Decisions on where to move stones in order to reach the river bed to avoid leakage from the Surber frame, have vital impact on results. The Surber is, in addition, limited to depths less than 40 cm. Traditionally, three to ten Surber samples are collected from each locality. This corresponds to 0.5-2 m² of substrate, depending on the size of the Surber sampler. Some results indicate that the number of samples should at least be doubled in order to obtain statistical significance, but our experience is that the number of parallel samples need to be much higher (Stark, 1993, and own unpublished data). Similarly, the area sampled by the kick net method is in practice impossible to standardize in distance (m), sampling time or sample volume, rendering abundance measures unreliable. In short, the number of sample parallels needed to gain statistically significant abundance measures is in practice unknown and varies in time and micro-space for the same location. Moving sampling site a few meters within a locality will likely yield different abundances (Callanan, Baars, & Kelly-Quinn, 2008). Total sorting of all organisms from all samples is insurmountable, so different forms of homogenization and subsampling have been practiced, adding another layer of uncertainty to the results. Even with advanced statistical analyses of large data sets, abundance numbers are still approximations (Bonada, Prat, Resh, & Statzner, 2006; Diserud & Aagaard, 2002; Doberstein, Karr, & Conquest, 2000; Engen, Aagaard, & Bongard, 2011; Stark, 1993; Vallania & Corigliano, 2007). Accurate abundance per area is only possible to determine for each Surber sample, by counting every organism. The probability that abundances in one Surber sample will be the same as in a replicate is extremely small. Adding replicates will consequently alter statistical results.

An important question is whether an exact abundance measure per area has any practical meaning at all in running waters. We argue that for biomonitoring in general, and for WFD, crude counting of taxa abundances as rare, common or dominating yields satisfactory results. We suggest reporting abundances of each taxa as approximate numbers per minute sample, because conspicuous alterations from expected numbers are valuable in detecting perturbations (Donohue, McGarrigle, & Mills, 2006). We have not tested rate of abundance approximation error from our sampling protocol, but even if it is as high as 30-40 %, it will still serve biomonitoring purposes. Finding one or three specimens of a species is of less importance than finding none.

Similarly, estimating the number of e.g. Chironomidae to be 500 or 700 is insignificant compared to, say, finding 10 specimens in a sample. For example, in Central Norway, undisturbed rivers normally have about 200-400 *Baetis rhodani* per 1-minute kick sample. This species is known to be sensitive to acidic conditions and can be almost or completely absent in rivers with a pH below about 5.5 (Raddum & Fjellheim, 1990). On the other hand, under eutrophic conditions, abundances of *B. rhodani* may reach thousands in a one-minute sample. In Norway, rivers often include rapid streams coming from upstream, pristine areas, so drift of species from these pristine areas is a notorious source of error. Most indices are based on detection, not numbers, like the ASPT/BMWP and Raddum index, which can result in error.

Subsampling is a trade-off between sorting efficiency, saving money and the quest for quality data. In this report, we present sorting equipment suitable for effectively handling large samples (**Figs 2 and 3**). The visual field produced by this combination of magnification and a sliding and stable boom arm give sufficient and effective vision to identify most genera and many species within EPT and other taxa. At the same time, the boom arm allows a rapid movement over substantial amounts of sample material in a large tray. Approximations of abundances are registered during the process, while attention is centred on finding species. If completely sorted in Petri dishes, the sample presented in **Appendix 3** might have taken two days to finish. Most likely, additional species might have been found in such a total sorting of all organisms, and this is not practically or economically feasible in biomonitoring projects. To minimize the significance of not registering every species, our index is only based on common species, predicted to be found even in small samples (Feeley et al., 2012). This makes the index more robust.

Based on data from Atna, South Central Norway, Bongard et al. (2011) modelled a sampling protocol focusing on species detection, illustrating a well-known and general logic principle in biology: Larger samples contain more species, and rate of detection of new species will go down with increasing number of organisms observed (M. L. McGarrigle & Lucey, 1983). This general principle applies regardless of river order, season, locality or any other variable introduced by sampling techniques (**Figure 4**).

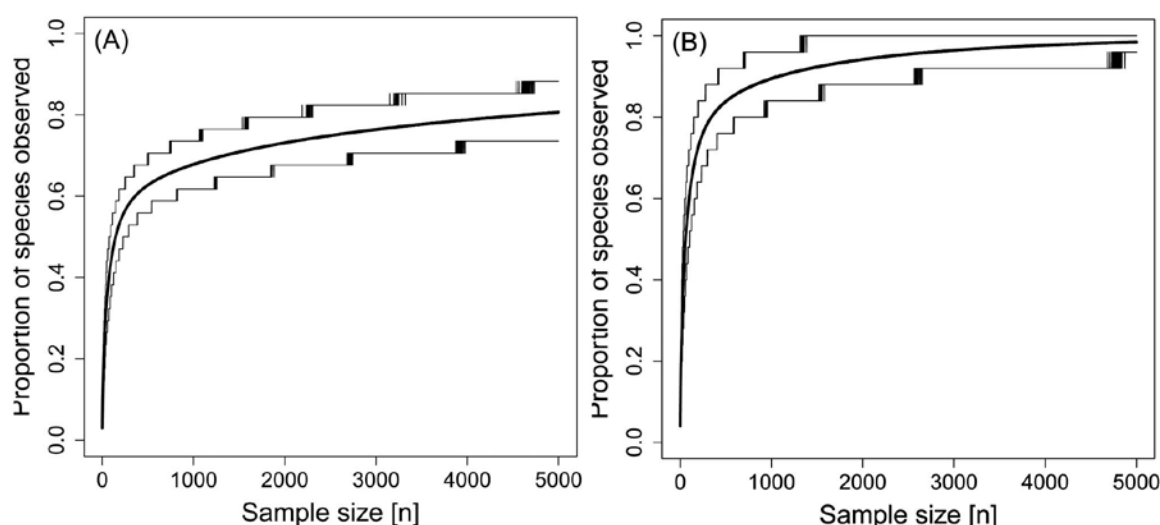


Figure 4. The proportion of species observed as a function of sample size for all 34 species (A) and for the 25 most common species only (B), simulated from the relative species abundance distribution for the disturbed community (solid lines). The thin lines indicate the 90% confidence belts, based on 10 000 simulated samples, data from Atna river, Central Norway (See T. Bongard et al., 2011 for statistical explanations).

Figure 4 shows a simulation for probability of detection: If a predicted species has a relative abundance of 0.001, it will require a sample size of around 3000 organisms to detect it with a probability greater than 0.95. Regardless of variances in sampling effort between areas and regions, sorting beyond a number of organisms at which new species become rare ($n \approx 3000$,

Figure 4), should provide a common species list of useful reliability. Bongard et al. (2011) suggested eight minutes of kicking to be enough for high probability sampling of the most common species expected to be present in Central Norway. In the present survey, we used 10 minutes of kicking.

Note that comparisons and classification boundaries are dependent on the total number of common species registered from the area. For Central Norway, more species are present than in the North-west, and ecological status class sizes will therefore be larger. Consequently, such conditions will provide a more accurate result. Data from North-West Norway are scarcer, so the expected EPT list will be shorter, and class boundaries therefore smaller. Accumulating data over time will contribute to a better and more extensive list and consequently adjust and improve class boundaries. This also illustrates that kick sampling time required to reach asymptotic levelling of new species detection varies considerably between regions or river types, in addition to the variables discussed earlier. The statistical detection model in **Figure 4** shows that too small a sample, whether being a limited number of minutes or meters of kicking, will result in a species detection placed at an unknown point on the steep part of the curve. A parallel sample will stochastically land on another point. To conclude, different sampling methods and their numerous weaknesses become less important with increasing sample size.

The above observations lay the ground for a standard based on diminishing species detection rate, regardless of sample size per se. When sorting larger samples one reaches the upper, flat part of the curve. This is determined by the fact that detection of new species becomes rare, and at this point there is a high probability that common species have already been registered. In the present survey, many samples would have reached this level much earlier than 10 minutes of sampling, but necessary sampling time cannot be known beforehand. Once sorting is complete, sample size threshold can then be evaluated. In practice, four to six minutes is often enough. Common species are usually easy to identify, while a relatively small number of species occupy a disproportionate amount of identification time. This represents a general problem throughout taxonomy as a science (Colwell & Coddington, 1994). Omitting common, but difficult taxa is one way of further increase efficiency. Preserving specimens of interesting taxa will secure data for the future.

We have focused on the common EPT species for the IBIBI index, but any invertebrate group may be used. A species list is nevertheless the ultimate basis to reveal environmental impacts. The IBIBI is also adaptable to only one sampling per year, just by lowering the benchmark, i.e. by adjusting expected species number to sampling date. In practice, benchmark EPT species lists will be a trade-off between variables like ecoregion size, river types, altitude, geography etc. Knowledge of these and other variables accumulates over time, but later changes in benchmark species lists will of course not change registered species lists from previous surveys. Insufficient sampling effort will provide uncertain information on the presence or absence of both common and uncommon species. Sample size must be very large before reaching a conclusion that red-listed or rare species are absent. The new technique of environmental DNA (eDNA) offers potential but requires development and calibration. Before eDNA becomes a reliable and practical method, extensive sampling remains the best method for registering low abundance species.

The boundaries between quality classes, i.e. the number of species deviating from expected pristine conditions, are based on the WFD description. The terms “High”, “Good”, “Moderate”, “Poor” and “Bad” represent qualitative values which are difficult to define exactly by a number. The WFD status classes are not exact descriptions of nature, but rather quality assessments open for discussion. Again, unambiguous sampling data consisting of species lists can be reassessed at any later moment in the light of new evaluation practice.

We have here classified the boundaries with equal numbers of species per class. The boundaries we suggest in **Table 2** have not yet been agreed as a benchmark in Norway, but these boundaries can easily be changed on the background of improved knowledge and experience. It is possible, for example, that the Good and Moderate classes should be larger, i.e. defined by more

species. Such changes can easily be adopted anytime. Data obtained from our sampling protocol will not lose its value by later adaptation in result interpretation. In addition, other criteria can be added that change the ecological status class, such as considering the number of rarer species, red-listed species etc. The species benchmark can readily be adjusted for a specific season. Sampling in autumn is preferred in Norway, as most species are present at this time. A general problem for indices is the fact that species numbers may fluctuate considerably between years. By focusing on common species, this problem is reduced, but one must still evaluate results on the background on yearly fluctuations.

Improvements of our preliminary benchmark lists can be made in several ways:

- Include species from other taxa. We have mentioned the class Oligochaeta. Ideally, Chironomidae should be considered one of the most important families in freshwater surveys, with hundreds of species and large abundances expected in all localities. Alas, species determination is extremely difficult and time-consuming, and very few experts with adequate competence are available. An eDNA approach has large potential for this family.
- Expand ecological status classes in accordance with the number of other EPT species registered. Such practice could be very useful, especially for a more accurate determination of the good-moderate boundary.
- Presence of red-listed species. Such registrations must always be subject to more detailed considerations.

4.1 ASPT shortcomings

The ASPT index has been adopted as the main index for classification of running water throughout Europe (Armitage, 1983). The index is very crude, only based on family taxa, and mostly adapted to the British Isles. It has long been known that species within the same family often have very different tolerance limits for metals, acidification or organic pollution (V. H. Resh & Rosenberg, 1993; V. H. Resh & Unzicker, 1975). Here, we discuss only a few examples of what we consider as shortcomings in the ASPT index. Experts within each family presented in the ASPT should be consulted for discussions on how to revise or change the index approach.

The class Oligochaeta in Norway consists of more than 50 species, ranging from very pollution tolerant to rather sensitive species (Bremnes & Storeid, 1994). Still, all species as a class are given the lowest score in the ASPT table. To include the class in the first place illustrates a fundamental problem with this form of index: Finding no Oligochaeta will increase the ASPT score by removing a 1 in the equation, even though a lack of these organisms may very well indicate large perturbations of many forms.

Most of the families within Ephemeroptera, Plecoptera and Trichoptera yield the highest ASPT score of 10. These families include Leptophlebiidae, Siphonuridae, Taeniopterygidae, Leuctridae, Chloroperlidae, Leptoceridae, Lepidostomatidae and Sericostomatidae. In Norway, many of these families have only one or two species with mostly wide distributions and abundances. In our experience we regard many of these species to be rather tolerant. For example, *Leptophlebia marginata*, *L. vespertina*, *Brachyptera risi*, *Siphonoperla burmeisteri*, *Lepidostoma hirtum* and *Sericostoma personatum* seem to be regularly found in streams considered to be polluted in Norwegian terms. There may be several reasons as to why this is so. The definition of pollution may in itself be unclear. Norwegian rivers are also generally fast-flowing, and the risk of catching specimens drifting from cleaner headwaters upstream is high. Particularly, the Baetidae family is very susceptible to this.

The trichopteran family Limnephilidae consists of several dozen species, of which some are extremely sensitive, like some *Apatania* spp. Others are tolerant, like many *Limnephilus* spp. The whole family is nevertheless given the number 7.

The mayfly *Ameletus inopinatus* is very common in Norway. It seems to tolerate fluctuating water levels in regulated rivers particularly well. It belongs to the family Ameletidae, which is not

included in the ASPT family list. The trichopteran family Glossosomatidae is also not included. It consists of several common and both sensitive and tolerant species. Several other families need to be discussed.

In **Table 6** we compare ASPT autumn score from the six rivers we surveyed. By omitting two of the taxa, Baetidae and Oligochaeta, the score rises considerably. *Baetis rhodani* is Norway's most common mayfly and is present in large numbers, except in acidic localities. The species has an ASPT score of only 4, and if missing increases the ASPT result towards a better ecological status. At the same time, if these two taxa are missing, it is likely that the locality is severely affected, and values should be lowered. The vulnerability of the index is clearly shown in this example. Indeed, introducing any sensitive species from the lower half of the BMWP/ASPT list will lower the final score.

Table 6. Comparisons of ASPT scores derived from autumn samples in the six rivers investigated. Lower line are values if no *Oligochaeta* or *Baetidae* were found.

	Søya	Ålvunda	Borgelva	Grana	Ogna	Verdals-elva
ASPT score	6,8	6,7	6,5	7,3	6,0	6,4
ASPT Score omitting Baetidae and Oligochaeta	7,5	7,4	6,9	7,7	6,6	7,1

4.2 Requirements for a robust and practical index

Bonada et al. (2006) identify twelve criteria for an ideal biomonitoring index for running water invertebrates. They classify these criteria in three categories. To sum up, we discuss each criterion in relation to the sampling protocol and the IBIBI index we propose.

Rationale:

1. Derived from sound theoretical concepts in biology.

Our proposal is a very simple Observed/Expected (O/E) approach firmly based on common species occurrences and approximate abundances, which are the ultimate measures of biodiversity and impact on an ecosystem. The simplicity strengthens the IBIBI, since it is based upon species presence or absence, collected in a way that reduces methodological problems.

2. A priori predictive.

A predicted species list is the basis for assessing community tolerance for all forms of perturbations. The calibration of the IBIBI to the five status classes (i.e. the number of species determining each status class) may change with time, but the data quality will not. A similar approach has been proposed by botanists by looking at species subsets as indicators of ecological status (Vellend, Lilley, & Starzomski, 2008). However, plant species are distributed patchily over larger areas. Freshwater invertebrates, and more so the common EPT species, normally have more continuous distributions, and are therefore more a priori predictive and suitable for O/E indices.

3. Potential to assess biological functions.

Biological functions are linked to species. Hence, species detection will constitute the basis of assessing biological functions. Functional groups are themselves constituted by species, and ecological functions in rivers are determined by the species present at any time. V. H. Resh and Rosenberg (1993) have stated that species is the most reliable level of taxonomy when it comes to indicator organisms for all forms of impacts.

4. Potential to discriminate overall human impact (i.e., to identify anthropogenic disturbance).

Perturbations of ecosystems are generally the outcome of human activities. Ecosystem deviations originating from natural causes are rare exceptions and normally easily identified as such. Again, the ultimate measure of any form of ecosystem impact is deviance from expected species biodiversity. Comparisons between undisturbed, intermediate and heavily polluted sites can be used to tune the scale in accordance with the WFD definitions. With increasing species knowledge, species prediction lists could easily include more of the rarer species. It is possible to merge ecoregions that have a similar fauna in order to establish shorter species lists valid for larger areas. With relatively little effort, an initial version of ecological reference conditions for every region in Norway can be presented. We believe that this also is possible for most European countries.

5. Potential to discriminate different types of human impact (i.e., to identify specific types of anthropogenic disturbance).

Macroinvertebrates like EPT are widely used in monitoring water quality as they exhibit a relatively wide range of responses to chemical and physical water quality stress (Karr & Chu, 1999; Simon, 2003). The effects of acidity, inorganic and organic pollution, and hydromorphological changes affecting biodiversity should be handled by an index in order to operationalize the WFD scale and make it a useful tool for managing freshwater.

All forms of influences on an ecosystem alter the distributions and abundances of species (Dobiesz et al., 2010). Each species might react differently to specific impacts, for example the response of *Baetis* spp. to acidity. Substantial deviances of species composition and abundances over time are rare in the absence of human impact. Regardless of perturbation or form of pollution, whether inorganic or organic, chemical or biological, deviation from expected biodiversity is the most unambiguous and ultimate measure of impact. Once any impact is established, the next step will be to analyze specific species reactions, and which forms of impact and concentration each species responds to (den Brink, Blake, Brock, & Maltby, 2006; Rubach, Baird, & Van den Brink, 2010; Von Der Ohe & Liess, 2004).

Implementation:

6. Low costs for sampling and sorting (field approaches) or for standardized experimentation (laboratory approaches).

The sampling protocol we propose substantially lowers laboratory time and expenses, by trading the quest for accurate numbers with an increase in effort to detect species. Hypothetically, it is possible to sample all species, including the rare ones, in any specific location (Vellend et al., 2008). However, for all practical purposes, surveys must strike a balance between available resources and the need for reliable knowledge.

7. Simple sampling protocol.

The strength of the IBIBI is the simplicity of both the sampling and the analysis of the results. No sophisticated homogenization or subsampling procedures are needed. By avoiding sorting and counting of large numbers of organisms, resources are automatically redirected to detection of red listed species. Bonada et al. (2006) conclude: “....a tool that could use a simple sampling protocol, i.e. least possible standardized techniques, such as 10 min of random kick sampling, collected once at any time of the year with a 0.5 mm mesh-sized net, would be advantageous for practical reasons in routine biomonitoring programs”. Sampling once a year is not enough to conclude on seasonal species composition, but our sampling protocol is cost effective and enables collection of multiple samples per year within the current economic costs. Variations in hatching time become less important when collecting more than one time per year. The index can, nevertheless, be calibrated to one sampling per year, by reducing the predicted species list to fit the specific season, and accordingly recalibrate the WFD ecological classification scale.

8. Low cost for taxa identifications (no specialists in taxonomy required).

The most common species are normally the simplest to identify. If not, difficult species may be omitted from the benchmark list, or taxa can be moved to a generic level. Another advantage from focusing on common species is that inter-annual species abundance variability become

less important. Common species tend to be at least present in detectable numbers every year. An argument for why some species are common may be that it is these that are least susceptible to various forms of environmental impact. Consequently, they are only expected to disappear in response to stronger environmental impacts. We believe this argument is overemphasized. For example, the most common mayfly in Norway, *Baetis rhodani*, is one of the species most sensitive to acid rain. In fact, of the 27 species suggested in our index, 14 are sensitive to acidity, according to Raddum's acidity index (Raddum & Fjellheim, 1990). Similar differences in sensitivity to various impacts might be present for each species. EPT species in general are sensitive and delicate organisms that react quickly to environmental alterations. A minority of EPT species are tolerant to environmental impacts, regardless of commonness. Regardless of differences in species sensitivity, a general decline in biodiversity is the ultimate and most robust measure of any form of perturbation.

Performance:

9. Large-scale applicability (across ecoregions or biogeographic provinces).

The IBIBI offers a solution to the problem of intercalibration. Each locality is compared to an Observed/Expected ratio of EPT species derived from the region or river class it locally belongs to. National experts can provide prediction lists of common species for their region and river types. Alas, for large parts of Europe, ecological reference conditions are lost forever. Nevertheless, sufficient species information from adjacent regions and/or historical material are usually at hand. Changes in knowledge of species distributions will easily lead to re-evaluation of species lists and the WFD class boundaries. A reliable sampling procedure providing unambiguous results will uphold its quality and value over time, independent of changes in ecoregions or river classes, which might lead to changes in expected species lists.

10. Reliable indication of changes in overall human impact.

The robustness of the species registration from our sampling protocol provides an ultimate and reliable indication of any deviance from a pristine ecosystem state. Alterations of biodiversity over time following changes in human activity are registered. To collect reliable data over time is a prerequisite for such evaluations. Shifts in abundances of species are natural and regularly occurring. Therefore, our sampling protocol include subsampling of approximate abundances, sufficient to evaluate large discrepancies in abundances of expected taxa.

11. Reliable indication of changes in different types of human impact.

Arguments from criterion 5 are valid also here. Different impacts affect different species in different ways. This knowledge can be used to describe and address types of perturbations on ecosystems. Reliability lies with the focus on species and their specific qualities. An interesting example from Norway is the fact that many rivers and streams are oligotrophic, and organic pollution consequently often increase both species distributions and abundances. Our sampling protocol will detect such changes.

12. Human impact indication on a linear scale.

The IBIBI and sampling protocol will detect minor changes in species registration over time and will provide data for evaluation of linear degradation. By accumulating knowledge of each species and its preferences, historical data of species deviances over the years sampled with the IBIBI method will constitute a background for evaluating human impact on a linear scale. Extreme localities to which species predictions do not apply are occasionally found. They could be identified as deviations from the expected values and described as such. WFD Annex II, no. 1.4 lists a number of situations that an index should be able to intercept, like point pollution, diffuse pollution, regulations, and drainage and morphological disturbances. These are all forms of stress that alter species composition and will be addressed by the IBIBI approach.

5 Conclusion

Comparing each area, region or ecotone to its own biodiversity benchmark, offers a solution to the intercalibration problem within the WFD. One way of achieving this is to increase sample size to ensure that common species are expected to be registered. Any observed deviation can be calibrated to the five ecological status classes of the WFD. Each region, ecotone or river type may have its own scale, and yet still be intercalibrated through the five ecological status classes. Species data from each locality is unambiguous and timeless, and results may continually be adjusted to the WFD index scale descriptions for each region as knowledge of expected species distributions and abundances increase over time. This way of thinking meets all the requirements presented by Bonada et al. (2006).

6 References

- Aagaard, K., Bækken, T., & Jonsson, B. (2002). *Felles instituttprogram. Virkninger av forurensninger på biologisk mangfold: Vann og vassdrag i by- og tettstednære områder. Sluttrapport 1997-2001* (19).
- Aagaard, K., & Dolmen, D. (1996). *Limnofauna Norvegica*: Tapir forlag.
- Aagaard, K., Solem, J. O., Bongard, T., & Hanssen, O. (2004). Studies of aquatic insects in the Atna river 1987-2002. In O. T. Sandlund & K. Aagaard (Eds.), *The Atna river: Studies in an Alpine-Boreal Watershed. Hydrobiologia* 521 (Vol. 521, pp. 87-105): Kluwer Academic Publ.
- Armitage, P. D., Moss, D., Wright, J. F., Furse, M. T. (1983). The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. *Water Res.*, 17(3), 333-347.
- Aroviita, J., Mykka, H., & Hamalainen, H. (2010). River bioassessment and the preservation of threatened species: Towards acceptable biological quality criteria. *Ecological Indicators*, 10(4), 789-795. doi:10.1016/j.ecolind.2009.12.007
- Arthington, A. H., Naiman, R. J., McClain, M. E., & Nilsson, C. (2010). Preserving the biodiversity and ecological services of rivers: new challenges and research opportunities. *Freshwater Biology*, 55(1), 1-16. doi:10.1111/j.1365-2427.2009.02340.x
- Bennett, C., Owen, R., Birk, S., Buffagni, A., Erba, S., Mengin, N., . . . Wasson, J. G. (2011). Bringing European river quality into line: an exercise to intercalibrate macro-invertebrate classification methods. *Hydrobiologia*, 667(1), 31-48. doi:10.1007/s10750-011-0635-2
- Bonada, N., Prat, N., Resh, V. H., & Statzner, B. (2006). Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annual Review of Entomology*, 51, 495-523.
- Bongard, T. (2005). *Effekter på bunndyr av aluminiumstilsetning mot G. salaris i Batnfjords-elva, 2003 og 2004*. (NINA Rapport 9). Retrieved from
- Bongard, T. (2009). *Effekter på bunndyr av rotenonbehandling mot G. salaris i Steinkjervassdragene 2009*.
- Bongard, T., & Aagaard, K. (2006). *BIOKLASS. Klassifisering av økologisk status i norske vannforekomster - elver. Forslag til bunndyrindeks for definisjon av Vanndirektivets fem nivåer for økologisk status* (NINA Rapport 113, 28 s.). Retrieved from
- Bongard, T., Diserud, O. H., Sandlund, O. T., & Aagaard, K. (2011). Detecting Invertebrate Species Change in Running Waters: An Approach Based on the Sufficient Sample Size Principle. *Benthic Open Environmental & Biological Monitoring Journal*, 4, 72-82.
- Bongard, T., Koksvik, J. I. (1989). *Lokal forurensning i Nidelva og en del tilløpsbekker vurdert på grunnlag av bunnfaunaen* (2).
- Bremnes, T., & Sloreid, S.-E. (1994). *Fåbørstemark i ferskvann. Utbredelse i Sør-Norge*. - *NINA Utredning* 56:1-42.
- Buffagni, A., Erba, S., & Furse, M. T. (2007). A simple procedure to harmonize class boundaries of assessment systems at the pan-European scale. *Environmental Science & Policy*, 10(7-8), 709-724. doi:10.1016/j.envsci.2007.03.005
- Buffagni, A., & Furse, M. (2006). Intercalibration and comparison - major results and conclusions from the STAR project. *Hydrobiologia*, 566, 357-364. doi:10.1007/s10750-006-0084-5
- Burdett, A. S., Fencl, J. S., & Turner, T. F. (2015). Evaluation of freshwater invertebrate sampling methods in a shallow aridland river (Rio Grande, New Mexico). *Aquatic Biology*, 23(2), 139-146. doi:10.3354/ab00616
- Callanan, M., Baars, J. R., & Kelly-Quinn, M. (2008). Critical influence of seasonal sampling on the ecological quality assessment of small headwater streams. *Hydrobiologia*, 610, 245-255. doi:10.1007/s10750-008-9439-4.
- Carter, J. L., & Resh, V. H. (2001). After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. *Journal of the North American Benthological Society*, 20(4), 658-682.

- Colwell, R. K., & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London B*, 345, 101-118.
- den Brink, P. J., Blake, N., Brock, T. C. M., & Maltby, L. (2006). Predictive value of species sensitivity distributions for effects of herbicides in freshwater ecosystems. *Human and Ecological Risk Assessment*, 12(4), 645-674. doi:10.1080/10807030500430559
- Directive 2000/60/EC of the European Parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities. L 327: 1-72., (2000).
- Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Rivers and lakes - Typology, reference conditions and classification systems. 10: 94 pp., (2003).
- Common implementation strategy for the Water Framework Directive (2000/60/EC). Guidance on the intercalibration process 2004-2006. 14: 31 pp., (2005).
- Diserud, O. H., & Aagaard, K. (2002). Testing for changes in community structure based on repeated sampling. *Ecology*, 83(8), 2271-2277.
- Doberstein, C. P., Karr, J. R., & Conquest, L. L. (2000). The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology*, 44(2), 355-371.
- Dobiesz, N. E., Hecky, R. E., Johnson, T. B., Sarvala, J., Dettmers, J. M., Lehtiniemi, M., . . . Witte, F. (2010). Metrics of ecosystem status for large aquatic systems - A global comparison. *Journal of Great Lakes Research*, 36(1), 123-138. doi:10.1016/j.jglr.2009.11.003
- Donohue, I., McGarrigle, M. L., & Mills, P. (2006). Linking catchment characteristics and water chemistry with the ecological status of Irish rivers. *Water Research*, 40(1), 91-98. doi:10.1016/j.watres.2005.10.027
- Engen, S., Aagaard, K., & Bongard, T. (2011). Disentangling the effects of heterogeneity, stochastic dynamics and sampling in a community of aquatic insects. *Ecological Modelling*, 222(8), 1387-1393. doi:10.1016/j.ecolmodel.2011.01.008
- Erba, S., Furse, M. T., Balestrini, R., Christodoulides, A., Ofenbock, T., van de Bund, W., . . . Buffagni, A. (2009). The validation of common European class boundaries for river benthic macroinvertebrates to facilitate the intercalibration process of the Water Framework Directive. *Hydrobiologia*, 633(1), 17-31. doi:10.1007/s10750-009-9873-y
- Official Journal of the European Union Commission decision of 30 October 2008 establishing, pursuant to Directive 2000/60/ EC of the European Parliament and of the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise. 332: 20-44., (2008).
- Feeley, H. B., Woods, M., Baars, J.-R., & Kelly-Quinn, M. (2012). Refining a kick sampling strategy for the bioassessment of benthic macroinvertebrates in headwater streams. *Hydrobiologia*, 683(1), 53-68. doi:10.1007/s10750-011-0940-9
- Frost, S., Huni, A., & Kershaw, W. E. (1971). Evaluation of a kicking technique for sampling stream bottom fauna. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 49, 167-173.
- Furse, M. T., Hering, D., Brabec, K., Buffagni, A., Sandin, L., & Verdonshot, P. F. M. (Eds.). (2006). *The Ecological Status of European Rivers - Evaluation and Intercalibration of assessment methods. Developments in Hydrobiology* (Vol. 566). Dordrecht, The Netherlands: Springer.
- Haase, P., Lohse, S., Pauls, S., Schindehutte, K., Sundermann, A., Rolaufts, P., & Hering, D. (2004). Assessing streams in Germany with benthic invertebrates: development of a practical standardised protocol for macro invertebrate sampling and sorting. *Limnologica*, 34(4), 349-365. doi:10.1016/s0075-9511(04)80005-7
- Haase, P., Murray-Bligh, J., Lohse, S., Pauls, S., Sundermann, A., Gunn, R., & Clarke, R. (2006). Assessing the impact of errors in sorting and identifying macroinvertebrate samples. *Hydrobiologia*, 566, 505-521. doi:10.1007/s10750-006-0075-6
- Haase, P., Pauls, S., Sundermann, A., & Zenker, A. (2004). Testing different sorting techniques in macro invertebrate samples from running waters. *Limnologica*, 34(4),

- 366-378. doi:10.1016/s0075-9511(04)80006-9
- Haase, P., Pauls, S. U., Engelhardt, C. H. M., & Sundermann, A. (2008). Effects of sampling microhabitats with low coverage within the STAR/AQEM macroinvertebrate sampling protocol on stream assessment. *Limnologia*, 38(1), 14-22. doi:10.1016/j.limno.2007.06.002
- Haase, P., Pauls, S. U., Schindehutte, K., & Sundermann, A. (2010). First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. *Journal of the North American Benthological Society*, 29(4), 1279-1291. doi:10.1899/09-183.1
- Hawkins, C. P., Norris, R. H., Hogue, J. N., & Feminella, J. W. (2000). Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications*, 10(5), 1456-1477.
- Karr, J. R., & Chu, E. W. (1999). *Restoring life in running waters: better biological monitoring*. Wash.: Island Press.
- McGarrigle, M. L., & Lucey, J. (1983). Biological monitoring of freshwaters. *Irish Journal of Environmental Science*, 2, 1-18.
- McGarrigle, M. L., & Lucey, J. (2009). Intercalibration of ecological status of rivers in Ireland for the purpose of the Water Framework Directive. *Biology and Environment-Proceedings of the Royal Irish Academy*, 109B(3), 237-246. doi:10.3318/bioe.2009.109.3.237
- MEA. (2005). *Millennium Ecosystem Assessment Synthesis Report* - Island Press. Retrieved from Washington DC.:
- Murray-Bligh, J., Buffagni, A., Cazzola, M., Birk, S., & Vlek, H. (2006). *Procedure for Calculating the ICMi Index: Step-By-Step Instructions For Central-Baltic river GIG. Version 2.6*. Retrieved from
- Nøst, T., Aagaard, K., Arnekleiv, J. V., Jensen, J. W., Koksvik, J. I., Solem, J. O. (1986). *Vassdragsreguleringer og invertebrater. En oversikt over kunnskapsnivået* (1).
- Ormerod, S. J., Dobson, M., Hildrew, A. G., & Townsend, C. R. (2010). Multiple stressors in freshwater ecosystems. *Freshwater Biology*, 55, 1-4. doi:10.1111/j.1365-2427.2009.02395.x
- Petrin, Z., Bækkelie, K. A. E., Bongard, T., Bremnes, T., Eriksen, T. E., Kjærstad, G., . . . Velle, G. (2016). *Innsamling og bearbeiding av bunndyrprøver – hva vi kan enes om*. NINA Rapport 1276. Retrieved from
- Raddum, G. G., & Fjellheim, A. (1990). Acid precipitation: Biological monitoring of streams and lakes. *The Science of the Total Environment*, 96, 57-66.
- Resh, V. H., & Rosenberg, D. M. (1993). *Freshwater biomonitoring and benthic macroinvertebrates*. New York: Chapman and Hall.
- Resh, V. H., & Unzicker, J. D. (1975). Water quality monitoring and aquatic organisms: The importance of species identification. *Journal WPCF*, 47(1), 9-19.
- Rubach, M. N., Baird, D. J., & Van den Brink, P. J. (2010). A new method for ranking mode-specific sensitivity of freshwater arthropods to insecticides and its relationship to biological traits. *Environmental Toxicology and Chemistry*, 29(2), 476-487. doi:10.1002/etc.55
- Simon, T. P. (2003). *Biological response signatures: indicator patterns using aquatic communities*. Boca Raton, Fla.: CRC Press.
- Stark, J. D. (1993). Performance of the Macroinvertebrate Community Index - Effects of Sampling Method, Sample Replication, Water Depth, Current Velocity, and Substratum on Index Values. *New Zealand Journal of Marine and Freshwater Research*, 27(4), 463-478.
- Storey, A. W., Edward, D. H. D., Gazey, P. (1991). Surber and kick sampling: a comparison for the assessment of macroinvertebrate community structure in streams of south-western Australia. *Hydrobiologia*, 211, 111-121.
- Ugedal, O., Berg, M., Bongard, T., Bremset, G., Kvingedal, E., Diserud, O. H., . . . Østborg, G. M. (2014). *Ferskvannsbiologiske undersøkelser i Surna. Sluttrapport for perioden 2009-2013*. <http://www.nina.no/archive/nina/PppBasePdf/rapport/2014/1051.pdf>
- Vallania, A., & Corigliano, M. D. (2007). The effect of regulation caused by a dam on the distribution of the functional feeding groups of the benthos in the sub basin of the

- Grande River (San Luis, Argentina). *Environmental Monitoring and Assessment*, 124(1-3), 201-209.
- Velle, G., Bækkeli, K. A., Arnekleiv, J. V., Bongard, T., Bremnes, T., Hall, J., . . . Stabell, T. (2018). *Kvalitetssikring av bunndyrbestemmelser i Norge. Uni Research Miljø LFI rapport 315. Uni Research Bergen. ISSN 1892-8889*. Retrieved from
- Vellend, M., Lilley, P. L., & Starzomski, B. M. (2008). Using subsets of species in biodiversity surveys. *Journal of Applied Ecology*, 45, 161-169.
- Vlek, H. E., Sporka, F., & Krno, I. (2006). Influence of macroinvertebrate sample size on bioassessment of stream. *Hydrobiologia*, 566, 523-542. doi:10.1007/s10750-006-0074-7
- Von Der Ohe, P. C., & Liess, M. (2004). Relative sensitivity distribution of aquatic invertebrates to organic and metal compounds. *Environmental Toxicology and Chemistry*, 23(1), 150-156.
- Walley, W. J., & Hawkes, H. A. (1996). A computer-based reappraisal of Biological Monitoring Working Party scores using data from the 1990 River Quality Survey of England and Wales. *Water Research*, 30(9), 2086-2094.
- Walley, W. J., & Hawkes, H. A. (1997). A computer-based development of the Biological Monitoring Working Party score system incorporating abundance rating, biotope type and indicator value. *Water Research*, 31(2), 201-210.

Appendix 1: UTM EU 89, zone 33 references for the localities around Trondheim:

1.

Homla	N 7033308 E 288570
Sagelva	N 7027667 E 282181
Nidelva	N 7025243 E 272496
Trollabekken	N 7043830 E 265706

Appendix 2: Example of freshwater invertebrates found in 10-minute kick samples from four localities around Trondheim. Sorting time presented at the bottom line.

Locality	Homla					Nidelva					Sagelva					Trollabekken				
Date	14.04.2016	25.05.2016	28.06.2016	06.09.2016	03.11.2016	12.04.2016	26.05.2016	28.06.2016	08.09.2016	03.11.2016	12.04.2016	25.05.2016	27.06.2016	06.09.2016	03.11.2016	14.04.2016	26.05.2016	28.06.2016	08.09.2016	03.11.2016
Turbellaria																1				
Crustacea																				
<i>Gammarus</i> sp.						400	400	2400	200	100										
Mollusca																				
<i>Radix balthica</i>	50	100		200	100	100	200	700	200	50								60		
<i>Gyraulus acronicus</i>					200	100	50		100	30								80		
Sphaeriidae		100			200	700	1300	1300	700			30								
Hirudinea																				
<i>Glossiphonia complanata</i>			1			50	10	50	50	50					10					
Oligochaeta	400	400	100	400	700	400	200	200	200	400	200	30	50	50	50	100	50		30	30
Hydrachnidae	50	200	100	100	200	400	100	100	200	50	30	50	100	50	50	30	50		30	50
Ephemeroptera																				
<i>Ephemera danica</i>				10			50	50	30			1		30						
<i>Ameletus inopinatus</i>		1				50	30	50			10	200								
<i>Baetis mulicus</i>	700	100	100					50					200			100	50	1300		
<i>B. niger</i>		50		2400	400		30						100	200			100			200
<i>B. rhodani</i>	2400	2400	400	4800	4800	400	2400	100	50	30	700	1300	700	700	700	400	2400	2400	400	400
<i>B. scambus</i>			100										700						200	
<i>Siphonurus lacustris</i>		1	30					50					50							
<i>Heptagenia</i> spp.					4800					700										
<i>H. dalecarlica</i>	200		400	200	700						10		1300	50	200					
<i>H. sulphurea</i>	200	200		200	700			200	50	400						30	100		200	200
<i>H. joemensis</i>			400										1300					200		
<i>H. fuscogrisea</i>				100					50											
<i>Ephemerella aroni</i>						100		700		200		10	200	30						
<i>E. mucronata</i>	1300	200	50		1300	2400	2400	200		30					200					
<i>Serratella ignita</i>			100						50					30						
<i>Leptophlebia</i> sp.			100						50	400				50						
Plecoptera																				
<i>Diura nanseni</i>	100	10	1	30	200		10		30		30		50	200	200	30		50	50	30
<i>Isoperla</i> spp.	100															100				
<i>I. difformis</i>				10																
<i>I. grammica</i>	50	100	10	50	200	1	30			100	30	100			100	30	30		50	30
<i>Siphonoperla burmeisteri</i>	50	30			100						100	100			50	100			30	30
<i>Taeniopteryx nebulosa</i>				50	50	30			50	100					100				30	50

<i>Brachyptera risi</i>	400	100								30	50			200	200	200			50
<i>Amphinemura borealis</i>	200	50		700		200	100	100		400	100	200		200					
<i>A. sulcicollis</i>	200	200	1		2400	10				100		100		200	100	200			
<i>Nemoura</i> sp.										200									
<i>N. avicularis</i>	30					10													
<i>N. cinerea</i>						1	50	100											30
<i>Nemurella pictetii</i>	50									200	10							30	30
<i>Protonemura meyeri</i>														200					
<i>Capnia</i> sp.(atra)						700				200				200					50
<i>C. pygmaea</i>											30								
<i>Leuctra</i> sp.																		50	
<i>L. fusca</i>				50					100					50				30	
<i>L. fusca/digitata</i>			1300			700		200				100	200				30		
<i>L. hippopus</i>	200	10		50	700		1			700	50			100	400	200			400
<i>L. nigra</i>														30				10	
Halipidae						10		50											
Scirtidae														10		30		10	10
Hydraenidae	50			50	100						100	100	200	200	50	50	100	100	100
Elmidae																			
<i>Elmis aenea</i>	50	30	50	50	100		50		50	50		200	50	30	50		400	50	
<i>Oulimnius tuberculatus</i>			30										30						
<i>Limnius volckmari</i>	10	400	50	50	100	30	10	100	30			400	50	30		1	30	50	
Sialidae																			
<i>Sialis</i> sp.									30										
<i>S. fuliginosa</i>							10	10	10		1					1			10
Värfluer																			
<i>Rhyacophila nubila</i>	50	100	50	50	100	30	30			30	1	30			50	50	10	10	50
<i>Aqapetus</i> spp.	100	100		200	1300	200	1300	200	50	700									
<i>Oxyethira</i> spp.														10				30	
<i>Hydroptila</i> spp.		10	30		100	1													
<i>Ithytrichia lamellaris</i>					50											10			
Philopotamidae						100													
<i>Philopotamus montanus</i>														30	30	30	1		50
<i>Wormalidia subnigra</i>																		10	
<i>Plectrocnemia conspersa</i>													10	10	30	10		30	10
<i>Polycentropus flavomaculatus</i>	30	30		30	50			30	30	50		30	30	30		30	100	30	30
<i>Neureclipsis bimaculata</i>						200	200	2400	400	700									
<i>Hydropsyche</i> spp.	100																		30
<i>H. pellucidula</i>	50	100	50	200	200					30	30						30		
<i>H. silfvenii</i>	50	100		200	400														
<i>H. nevae</i>						4800			30										
<i>H. siltalai</i>		30														30	50		
<i>Arctopsyche ladogensis</i>						1													
<i>Lepidostoma hirtum</i>			30	50	50	100	10		30	50		30							
Limnephilidae		200						50		10	10	50	100	50				30	
<i>Chaetopteryx villosa</i>										1									
<i>Halesus radiatus</i>	10		10					30										10	
<i>H. digitatus</i>		1																	

<i>Potamophylax cingulatus</i>						1						10				10				
<i>P. latipennis</i>		1						10								1				
<i>Apatania</i> spp.									30											
<i>A. stigmatella</i>						200														
<i>Silo pallipes</i>	10	1		50	30											1				
<i>Sericostoma personatum</i>				100			10	400		10	1	10	1	50	10	30	30	50	30	10
<i>Athripsodes</i> spp.	50	10	50	50	50		200	50	30	30										
<i>Ceraclea</i> spp.									50											
<i>C. annulicornis</i>						10				10										
Diptera indet.	30	30									30									
Tipulidae	100	100	50	50	200		100	50	30	50	50	30	30	10	30	200	100	100	30	30
Simuliidae	100	2400	100		700	100			50		200	1300			200	2400	700	400	50	100
Chironomidae	200	700	200	200	1300	2400	100	400	100	400	200	400	400	50	700	400	100	200	100	100
Ceratopogonidae			30			30	10		10		10				30					10
<i>Pericoma</i> sp.																30				
Total in 1 min kick sample	767	860	393	1068	1778	1497	938	1033	304	657	196	489	585	208	424	474	483	518	159	202
Sorting time, hours	1	2	1,5	1,5	2	2	2	2	1	1	0,5	1,5	1	1	1	1,5	1	1,5	1	1
Species determination time, minutes	20	20	20	20	30	30	30	30	30	30	15	15	15	15	20	15	15	15	15	15

Appendix 3. Example of sorting and subsampling of one sample with notes of abundances and species for each tray. This sample needed to be split into ten trays in total.

Locality, Date: Nidelva 06-28-2016											
Tray No.	1	2	3	4	5	6	7	8	9	10	TOTAL
Gammaridae	1000	500	200	100	200	200	100	30			2400
Lymnaeidae	10			10		10	10	150	200	300	700
Sphaeriidae		20	20	20	30	30	20	200	200	800	1300
Hirudinea		10	5	5		10	10	10			50
Oligochaeta	10	10	30	10		30	30	30	20	20	200
Hydrachnidae	10	30	5	20	5	10	5		5		100
Ephemeroptera											
<i>Ameletus inopinatus</i>	20	20		10							50
<i>Baetis muticus/niger</i>	10	10	20	10							50
<i>Baetis rhodani/sp</i>	30	50	20								100
<i>Ephemera sp</i>						20	20	10			50
<i>Heptagenia sp</i>	30	20	20	50	20	30	10	10			200
<i>Siphonuridae</i>	5	5	20		20						50
<i>Ephemerella aroni</i>	50	50	100	200	50	30	100	50	50		700
<i>Ephemerella mucronata</i>	30	10	30	50	20	20	30	10			200
Plecoptera											
<i>Amphinemura borealis</i>		10		10	50	20		10			100
<i>Nemoura sp.</i>	20	10	10	5	20	20	10	10			100
<i>Leuctra sp.</i>		20	20	20	20	20	30	20	30	20	200
Halipilidae			5	5	5	5	10	10	5	5	50
Elmidae											
<i>Limnius volckmari</i>		5	5	5	5	10	20	20	30	10	100
Sialidae				1		1	1		2	2	10
Trichoptera											
<i>Glossosomatidae</i>						50	20	10	100	20	200
<i>Polycentropus flavomaculatus</i>		10		10	10						30
<i>Polycentropodidae sp</i>	200	300	300	100	500	800	100	50	50		2400
Limnephilidae								10	20	20	50
<i>Halesus/Potamophylax/Chaetopteryx</i>			1	1		2	3	15	15	5	40
<i>Leptoceridae sp.</i>		3		5	5	5	5	10	10		50
<i>Sericostoma personatum</i>							200	100	200	50	400
Diptera											
Tipulidae			3		10	5	5	10	10	10	50
Chironomidae	50	30	50	50	30	100	30	30	10	10	400
Sorting time											2 h

The Norwegian Institute for Nature Research, NINA, is as an independent foundation focusing on environmental research, emphasizing the interaction between human society, natural resources and biodiversity.

NINA was established in 1988. The headquarters are located in Trondheim, with branches in Tromsø, Lillehammer, Bergen and Oslo. In addition, NINA owns and runs the aquatic research station for wild fish at lms in Rogaland and the arctic fox breeding center at Oppdal.

NINA's activities include research, environmental impact assessments, environmental monitoring, counselling and evaluation. NINA's scientists come from a wide range of disciplinary backgrounds that include biologists, geographers, geneticists, social scientists, sociologists and more. We have a broad-based expertise on the genetic, population, species, ecosystem and landscape level, in terrestrial, freshwater and coastal marine ecosystems.

ISSN: 1504-3312
ISBN: 978-82-426-3286-9

Norwegian Institute for Nature Research

NINA head office

Postal address: P.O. Box 5685 Torgarden,
NO-7485 Trondheim, NORWAY

Visiting address: Høgskoleringen 9, 7034 Trondheim

Phone: +47 73 80 14 00

E-mail: firmapost@nina.no

Organization Number: 9500 37 687

<http://www.nina.no>



Cooperation and expertise for a sustainable future