NIVA REPORT 7927-2024

Norsk institutt for vannforskning

ICP Waters Report 155/2024

Biological intercalibration: Invertebrates 2023

Photo: Christian Lucin Bodin (NORCE)

International Cooperative Programme on Assessment and Monitoring Effects of Air Pollution on Rivers and Lakes

Convention on Long-Range Transboundary Air Pollution

UNECE

Convention on Long-range Transboundary Air Pollution

Report

Serial no: 7927-2024

ICP Waters Report
155/2024

ISBN 978-82-577-7663-3
NIVA report
ISSN 1894-7948

This report has been quality assured according to NIVA's quality system and has been approved by:

Kari Austnes
Project Manager

Johnny Håll
Quality Assurer

Hans Fredrik Veiteberg Braaten
Research Manager

www.niva.no

Norwegian Institute for Water Research

Title	Pages	Date
Biological intercalibration: Invertebrates 2023	22 + appendix	16.01.2024

Author(s)
Christian Lucien Bodin[1]
Gaute Velle[1,2]
Torunn Svanevik Landås[1]

[1]NORCE Norwegian Research Centre AS
[2]University of Bergen, Bergen, Norway

Topic group
Monitoring

Distribution
Open

Client(s)	Client's contact person
Norwegian Ministry of Climate and Environment United Nations Economic Commission for Europe (UNECE)	Eli Marie Åsen

Published by NIVA
10300

Abstract
The ICP Waters biological intercalibration of invertebrates was executed to harmonise taxonomic work across countries and is of high value in programmes where the focus is on community analyses, e.g., for the classification of ecological status according to the EU Water Framework Directive.
The 27th biological intercalibration of invertebrates in ICP Waters included four participants. A total 99.9 % of the species and 99.9 % of the genera were correctly identified in 2023. The mean Quality assurance index (Qi) ranged from 98.1 to 100.0. The results show that the average Qi has remained above 80% since 1992, suggesting skilled taxonomists in the laboratories affiliated to ICP Waters.

Keywords: ICP Waters, EU Water Framework Directive, Aquatic fauna, Monitoring
Emneord: ICP Waters, EUs vanndirektiv, Akvatisk fauna, Overvåking

CONVENTION ON LONG-RANGE TRANSBOUNDARY AIR POLLUTION

INTERNATIONAL COOPERATIVE PROGRAMME ON ASSESSMENT AND MONITORING OF THE EFFECTS OF AIR POLLUTION ON RIVERS AND LAKES

Biological intercalibration: Invertebrates 2023

Prepared at the ICP Waters Programme Subcentre

NORCE AS

Bergen, January 2024

Table of contents

Preface

The International Cooperative Programme on Assessment and Monitoring of the Effects of Air Pollution on Rivers and Lakes (ICP Waters) was established under the Executive Body of the UNECE Convention on Long-range Transboundary Air Pollution (CLRTAP) in July 1985. Since then, ICP Waters has been an important contributor to documenting the effects of implementing the Protocols under the Convention. ICP Waters has prepared numerous assessments, reports and publications that address the effects of long-range transported air pollution.

ICP Waters and its Programme Centre are chaired and hosted by the Norwegian Institute for Water Research (NIVA), respectively. A programme subcentre is established at NORCE, Bergen. ICP Waters is supported financially by the Norwegian Ministry of Climate and Environment and the Trust Fund of the UNECE LRTAP Convention.

The main aim of the ICP Waters programme is to assess, on a regional basis, the degree and geographical extent of the impact of atmospheric pollution, in particular acidification, on surface waters. More than 20 countries in Europe and North America participate in the programme on a regular basis.

One objective of the ICP Waters programme is to establish and maintain an international network of surface water monitoring sites and promote international harmonization of monitoring practices. Inter-laboratory quality assurance tests are a tool in this work. Here any biases between analyses carried out by individual participants in the programme are identified and controlled. The tests are also a valuable tool for taxonomic discussions and the exchange of identification keys among the participating laboratories, thereby improving taxonomic skills.

Here we report the results from the 27th intercalibration of invertebrate fauna. We also compare results from all 27 intercalibrations.

Bergen, January 2024

Gaute Velle

ICP Waters Programme Subcentre

Summary

The ICP Waters biological intercalibration of invertebrates is important for harmonizing taxonomic work across countries and is of high value in monitoring programmes where the focus is on community analyses, e.g., for the classification of ecological status according to the EU Water Framework Directive. Intercalibration practices ensures high quality data in the ICP Waters database and increases the taxonomic skills of the participants. The intercalibration under the ICP Waters programme has run annually since 1992. It was the first regular intercalibration test for identification at species level in Europe. Here, we present results from the intercalibration in 2023 and trends in results from the intercalibration from the initial intercalibration in 1992 and up to the present.

The 27th biological intercalibration of invertebrates in ICP Waters included four participants. A total 99.9 % of the species and 99.9 % of the genera were correctly identified in 2023. Combined, the four laboratories misidentified only three individuals at species level and one at genus level (out of 276 specimen). The mean Quality assurance index (Qi) ranged from 98.1 to 100.0, where 80 is the limit for good taxonomic work. The highest mean Qi- score for the intercalibration in 2023 was for the Plecoptera group with 100.0, while the lowest mean score was from the miscellaneous taxa with 98.1. This deviates from a trend seen the last 26 years, where participants acquire highest scores for Trichoptera and the lowest score for Plecoptera. The average number of participating laboratories over time is 4.5. The results show that the average Qi has remained above 80% since 1992, suggesting skilled taxonomists in the laboratories affiliated to ICP Waters.

Sammendrag

ICP Waters' biologiske interkalibrering av virvelløse dyr er viktig for å harmonisere taksonomisk arbeid på tvers av land og har høy verdi i overvåkingsprogrammer der fokuset er på samfunnsanalyser, f.eks. for klassifisering av økologisk tilstand i henhold til EUs vanndirektiv. Interkalibreringspraksis sikrer høy kvalitet på dataene i ICP Waters-databasen og øker deltakernes taksonomiske ferdigheter. Interkalibreringen under ICP Waters-programmet har blitt gjennomført årlig siden 1992. Dette var den første regelmessige interkalibreringstesten for identifikasjon på artsnivå i Europa. Her presenterer vi resultater fra interkalibreringen i 2023 og trender i resultatene fra den første interkalibreringen i 1992 og frem til i dag.

Den 27. biologiske interkalibreringen av virvelløse dyr i ICP Waters inkluderte fire deltakere. Totalt sett ble 99,9 % av artene og 99,9 % av slektene riktig identifisert i 2023. De fire laboratoriene feilidentifiserte bare tre individer på artsnivå og ett på slektsnivå (av 276 individer). Gjennomsnittlig kvalitetssikringsindeks (Qi) varierte fra 98,1 til 100,0, der 80 er grensen for godt taksonomisk arbeid. Den høyeste gjennomsnittlige Qi-scoren for interkalibreringen i 2023 var for gruppen Plecoptera (steinfluer) med 100,0, mens den laveste gjennomsnittlige poengsummen var fra diverse taksoner med 98,1. Dette avviker fra en trend sett de siste 26 årene, der deltakerne oppnår høyest poeng for Trichoptera (vårfluer) og lavest poeng for Plecoptera (steinfluer). Gjennomsnittlig antall deltakende laboratorier over tid er 4,5. Resultatene viser at gjennomsnittlig Qi har vært over 80 % siden 1992, noe som tyder på dyktige taksonomer i laboratoriene tilknyttet ICP Waters.

1 Introduction

The purpose of the biological intercalibration of invertebrates is to evaluate the quality of the biological data delivered to the Programme centre. The data are used nationally and by ICP Waters to indicate environmental conditions from the species and their tolerances to acidification and other stressors (Raddum et al. 1988, Fjellheim and Raddum 1990, Raddum 1999, Velle et al. 2013, 2016). The significance of potential trends in biotic indices, both for a specific site/watershed and for comparisons of trends among regions or among countries, can be evaluated once the data quality is known. The data are also used in numerical analyses (Larsen *et al.* 1996, Skjelkvåle *et al.* 2000, Halvorsen *et al.* 2002, Halvorsen *et al.* 2003), and in analyses of biodiversity (Velle *et al*., 2013, Velle *et al*. 2016). The results from such data analyses are especially sensitive to the quality of the species identifications. The biological intercalibration focuses on taxonomic skills of the participants and is a tool for improving the quality of work at the different laboratories, as well as harmonization of the biological database.

The methods for biological intercalibration that we use, were outlined in 1991 at the seventh ICP Waters Task Force meeting in Galway, Ireland. The countries/laboratories should know their native fauna. Since fauna vary according to geographical regions, specific samples based on their native fauna are prepared for each participating laboratory. We cannot use standardized samples for all participants. Therefore, each laboratory sends identified samples of invertebrates from their own monitoring sites to the organizer (the Programme subcentre). The organizer adds species previously sampled and identified by the specific laboratory. Each laboratory receives individual test samples with species representing their own monitoring region. Each participant is therefore tested on their ability to identify fauna that is familiar to them. An important implication of this procedure is that the participant prepares the solution of the test, and that the organizer remains neutral without the ability to influence the results. To highlight that the organizer has little opportunity to influence the results, each participant is given the opportunity to comment on the results and agree on the conclusion from their part of the intercalibration.

The taxonomic skill of the participants is measured by using a quality assurance index (Raddum 2005). This index evaluates the skill of participants when identifying species and genera. It also considers the effort of identifying all specimens in the sample. The highest index score is 100, while a value of 80 is set as the limit of good taxonomic work.

This report mostly adheres to a similar format that has been used in previous reports and contains text partially or completely retained from previous reports (Raddum 2005, Fjellheim et al. 2014, Halvorsen et al. 2016, Velle et al. 2018).

2 Methods

Preparation of the test-samples

Samples of invertebrates were sent from all participating laboratories to the organizer at the ICP Waters subcentre. These samples were used to compose test samples, with the addition of specimens from earlier exercises and from collections at the subcentre. The test samples included caddis flies (Trichoptera), stone flies (Plecoptera), mayflies (Ephemeroptera) and miscellaneous.

Miscellaneous included water beetles (Coleoptera), crustaceans (Malacostraca), leeches (Hirudinea), molluscs (Gastropoda & Bivalvia), dragonflies (Odonata), water boatmen (Corixidae), midges and flies (Diptera), butterflies and moths (Lepidoptera) and true bugs (Heteroptera). Both larvae and adults were included. Leeches, molluscs, and crustaceans are sensitive to acid water and important for the evaluation of acidification. The tolerance of some miscellaneous species is poorly known. They are often regarded as tolerant to acidic water and of low importance for the evaluation of acidity. They are still important in invertebrate community analysis.

The geographical distribution of the taxa was checked using the Fauna Europaea Web Service 2013 (http://www.faunaeur.org). This is a database of the scientific names and distribution of multicellular European land and fresh-water animals (see example in **Figure 1**).

Identification

To minimize possible faults, the following procedure was used in preparing the test samples:

- The participating laboratory first identified the source material for the test samples and shipped the specimens to the organizer.
- Two persons from the organizing institution verified the identification of the specimen as far as possible without damaging the individuals.
- The content of two test samples per participant was listed in a table. Two persons controlled that the correct numbers and species were placed in the test samples according to the table.

Damage to the material

The quality of the test material may be diminished during handling and shipping. Taxonomically important parts of the body, such as gills, legs, cerci and mouthparts can be lost or damaged during identification, handling and transportation. Mixing of individuals between samples may occur during identification. All above mentioned examples are source of errors that could influence the process of identification and verification of taxa negatively, and thereby the end results.

Evaluation

The participants were invited to comment on the results before the report was published. In this way, we removed potential bias - for example misidentification caused by damaged test material. In cases of disagreement between the participant and the organizer, the material may be checked again by the organizer and by the participant. This procedure may act educational for both parts and ensures that both the participant and the organizer agree on the conclusions from the intercalibration.

Figure 1. *Geographical distribution of the caddisfly Rhyacophila nubila in Europe. This species is widely distributed but is absent from several West-European countries. Map after Fauna Europaea Web Service, http://www.faunaeur.org, Illustration: Arne Fjellheim*

For calculation of errors, we considered possible degradation of the material. Further, a misidentified species counted as only one fault, even if the sample includes many individuals of the species. We encouraged participants to give comments on matters that may impede the identification. For example, a misidentification will not count as a fault if a specimen lacks important taxonomic characters. Such comments must be made before the results are sent to the organizer. We have discriminated between short-comings in identification due to damaged material, and true errors (wrong species – or genus).

The organizer also noted how many specimens a participant identified per sample. This is referred to as *percent identified*. A low percent means that many individuals were not identified and will consequently reduce the value of the taxonomic work. In cases where more specimens were identified than sent to the laboratories, each excess specimen counted as one error.

Available material for making test samples vary. Normally, each laboratory receives between 60 and 130 species in the two samples. Samples with low diversity are easier to handle than samples with high diversity (see **Appendix B**). Handling time should therefore be kept in mind when the results are evaluated. Small samples were avoided, as only a few misidentifications could result in a low score.

There are 1814 species of European mayfly (Ephemeroptera), stonefly (Plecoptera) and caddisfly (Trichoptera) (http://www.faunaeur.org). However, the biodiversity differs between countries. Generally, the number of species decreases along a gradient from Southern to Northern Europe. This is also a fact to bear in mind when judging taxonomical capacity. As an example of this, the freshwater fauna of Switzerland is much richer than in Norway and Sweden – despite the fact that the area of Switzerland is approximately 1/10 of the two Nordic countries (**Figure 2**).

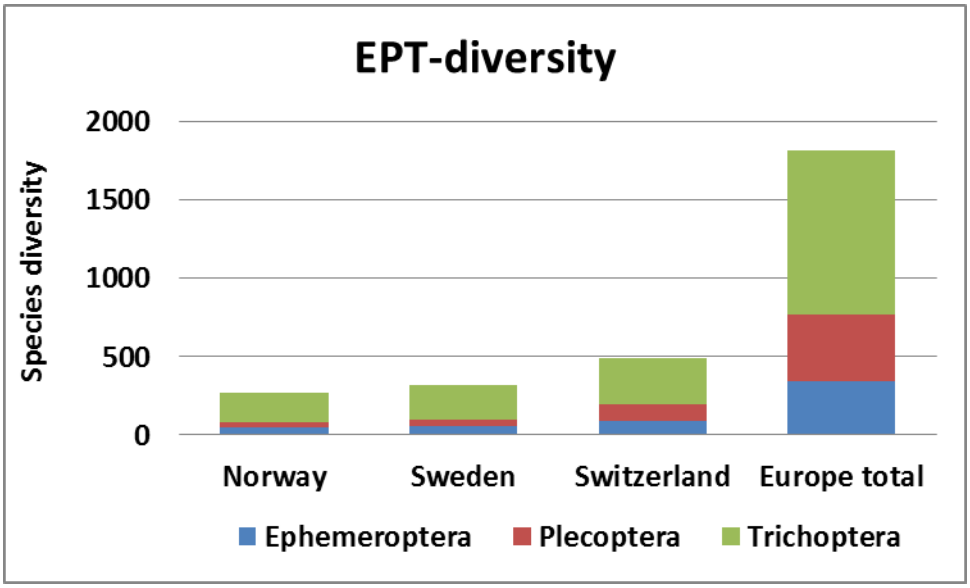


Figure 2. Species diversity of mayflies (Ephemeroptera), stoneflies (Plecoptera) and caddisflies (Trichoptera) in Norway, Sweden and Switzerland (after Fauna Europaea Web Service, http://www.faunaeur.org.

Quality assurance index

We have calculated the Quality assurance index, Qi, for invertebrate groups as well as the mean index for each participant. The Qi integrates the separate levels of the identifications as follows:

Qi = (% correct species/10) * (% correct genus/10) * (% identified individuals/100)

Qi will be a number between 0 and 100 with increasing skill. A score ≥ 80 is regarded as good and thus acceptable taxonomical work.

Test of the subcentre

The ICP Waters subcentre in Bergen is tested with the help from the Swedish participant every second year. The Swedish University of Agricultural Sciences in Uppsala prepares and evaluates the test taken from the subcentre. Methodology and implementation are otherwise identical to the other tests.

3 Results and discussion

Four laboratories participated in the intercalibration of invertebrates in 2023 (**Appendix A**). The species lists and the identification results are shown in **Appendix B**, **Tables 1-4**.

Mayflies (Ephemeroptera)

For the identification of mayflies (**Figure 3**), only one mistake occurred, where laboratory 3 misidentified one individual at species level. The lab still scored well above the limit for good taxonomic work.

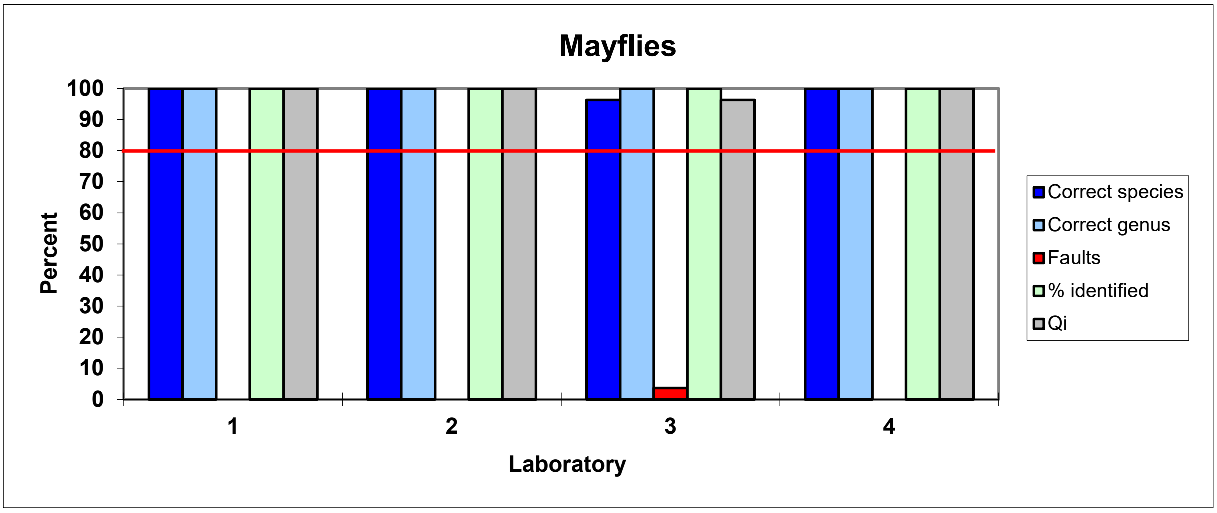


Figure 3. Results from the identification of mayflies. The red line indicates the limit for good taxonomic work. Qi = quality assurance index.

Stoneflies (Plecoptera)

Results for the identification of stoneflies are shown in **Figure 4**. All laboratories were flawless and identified all individuals correctly down to species level.

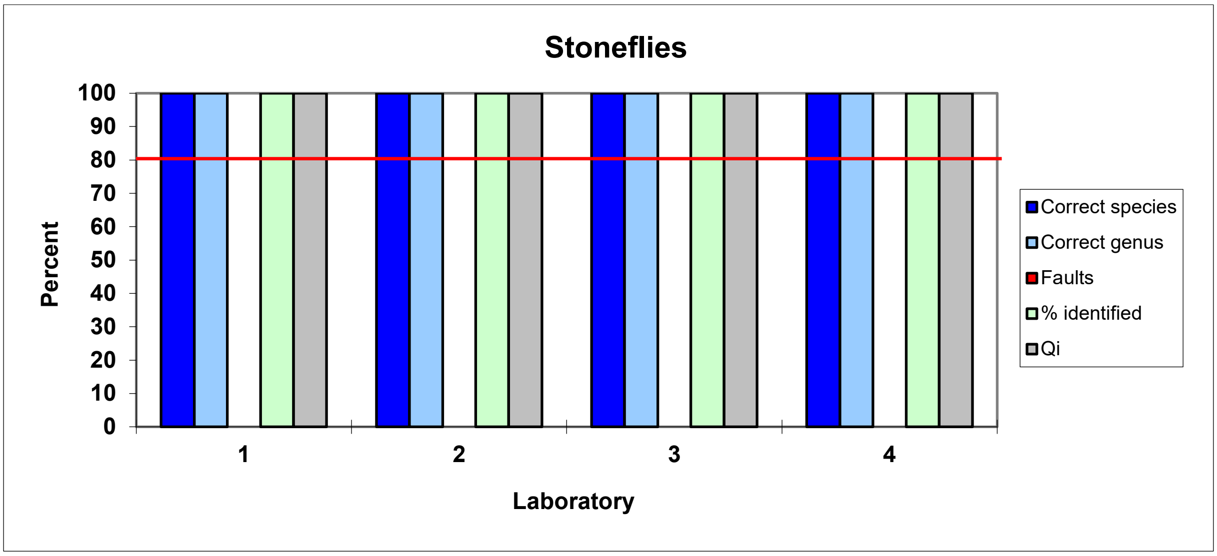


Figure 4. Results from the identification of stoneflies. The red line indicates the limit for good taxonomic work. Qi = quality assurance index.

Caddisflies (Trichoptera)

Laboratory 4 misidentified 1 individual at species level and 2 individuals of the same species at species level, receiving one error. Laboratory 4 still acquired a score above the acceptable limit of good taxonomic work (**Figure 5**). Laboratory 1, 2 and 3 identified all individuals of caddisflies correctly and receives a score of 100.

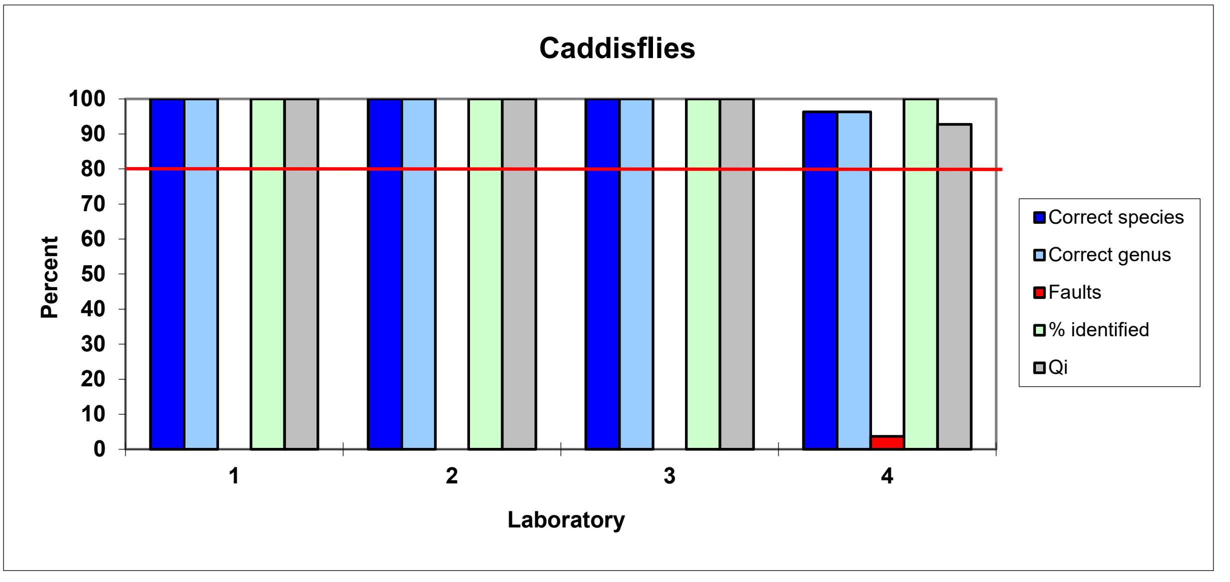


Figure 5. *Results from the identification of caddisflies. The red line indicates the limit for good taxonomic work. Qi = quality assurance index.*

Miscellaneous

Laboratory 4 misidentified 1 individual at species level and acquired a Qi score of 92.3 (**Figure 6**). This is still above the limit for good taxonomic work. Laboratory 1, 2 and 3 correctly identified all species in their samples and acquired a Qi score of 100.

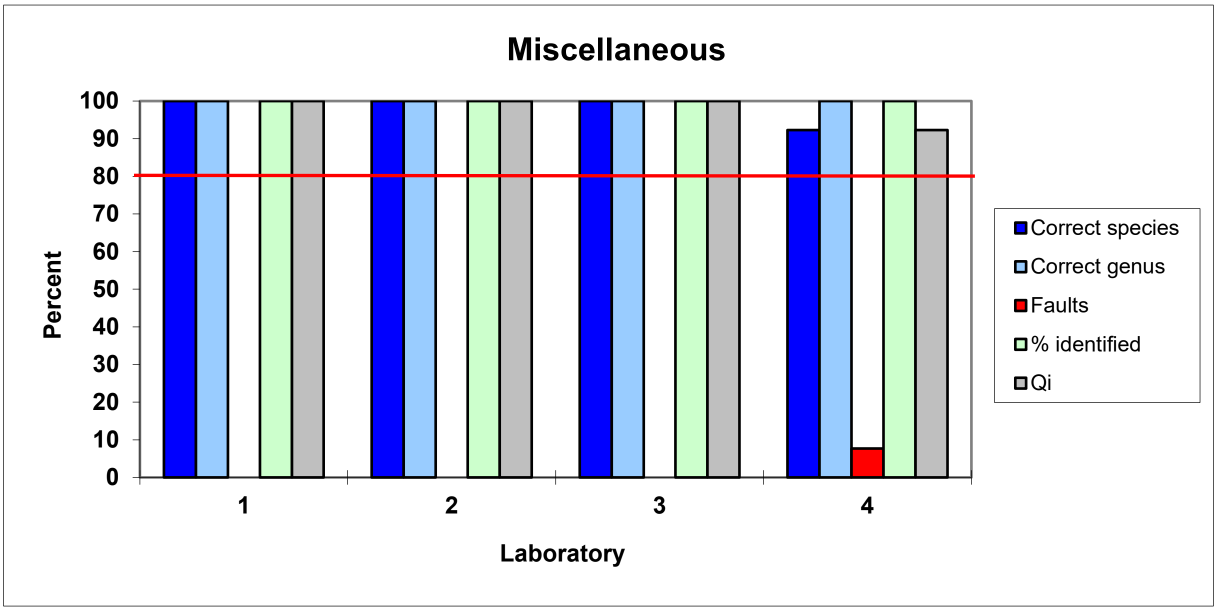


Figure 6. Results from the identification of miscellaneous groups of invertebrates. The red line indicates the limit for good taxonomic work. Qi = quality assurance index.

Total number of species in the sample

A total of 276 individuals were sent to the laboratories. Laboratory 1 received a total of 63 individuals, laboratory 2 received 53 individuals, laboratory 3 received 84 individuals and laboratory 4 received 76 individuals. All individuals were reported to the organizer.

4 Overall evaluation

The laboratories correctly identified a high proportion of the total number of species in the test and acquired a mean score higher than the average from the previous 26 intercalibration tests. The results for this year's intercalibration test got the highest score achieved since the intercalibration started in 1992. The four labs combined received a QI score of 99.0, 6.6 points above the average from all previous tests.

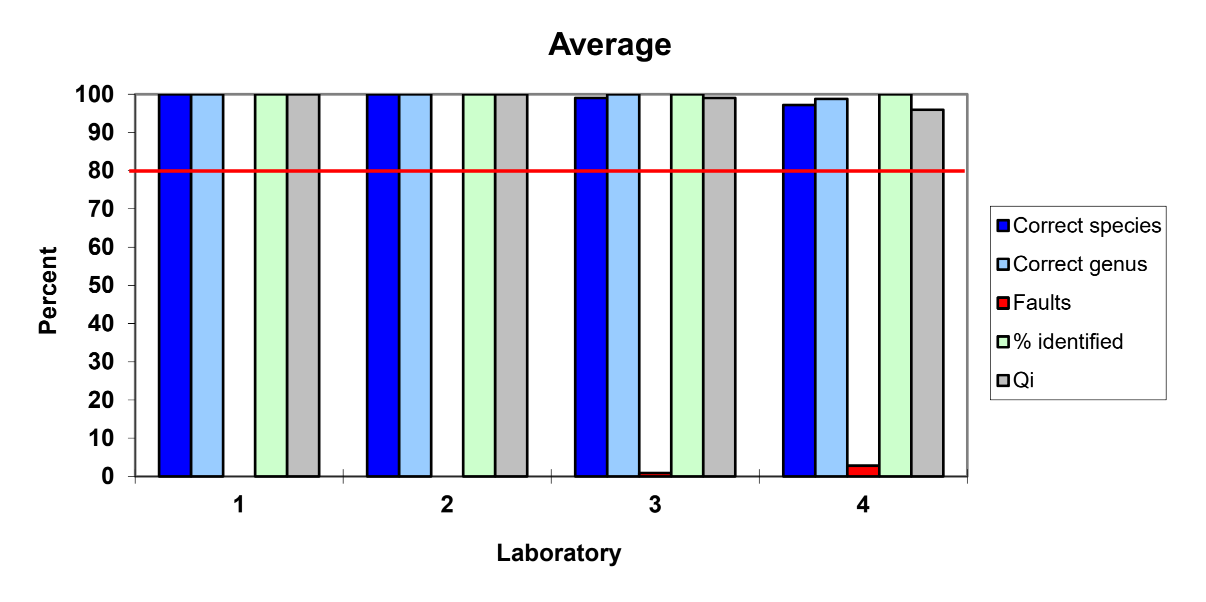


Figure 7. Mean skill in percent of identifying species and genus, and mean Qi for each laboratory. The red line indicates the acceptable limit. Qi = quality assurance index.

The highest mean Qi- score for the intercalibration in 2023 was in the group of Plecoptera with a score of 100%, while the lowest mean score was from the Caddisflies group with 95.6%. This deviates from a trend seen the past 26 intercalibration tests, where participants normally show highest skills in identifying Trichoptera and the lowest skills in identifying Plecoptera. This year, two laboratories achieved a 100% QI score in all categories.

The biological intercalibration is important for harmonizing biological material/databases and will be of high value in projects that focus on community analyses, or where the ecological status of waterbodies should be determined. The biological intercalibration under the ICP Waters programme was the first regular test aiming to test taxonomic skills in identifying benthic invertebrates. Today, similar tests are run by the North American Benthological Society[1] and by the Natural History Museum, London (Identification Qualifications – IdQ test). The invertebrate groups covered in the latter test are those used in the BMWP water quality score system (Armitage et al., 1983) and include groups used for monitoring freshwater environments under the EU water framework directive (Schartau et al. 2008). In 2018 and in 2020, NORCE also organized biological intercalibrations for Norwegian laboratories that identify benthic invertebrates on a regular basis. The result from the Norwegian tests indicated that participants assigned specimens from an identical sample to a significant different number of taxa and with a significantly different species composition (Velle et al. 2018, Velle et al. 2020). The differences resulted in a classification of ecological status that to some extent was person dependent (Velle et al. 2018). These results highlight the importance of quality assurance and coordination of species identifications. Because of these findings from the intercalibration in Norway, regular intercalibrations will be performed in the

[1] https://ncse.ngo/north-american-benthological-society

future. Also, the Norwegian Environment Agency use participations in intercalibrations as part of the evaluation criteria when assigning companies to new projects (Velle et al. 2020).

5 Trends over time

The invertebrate intercalibration in ICP Waters started in 1992. An overall high of 11 laboratories participated during the first intercalibration (**Figure 8**). Since then, the average has been just under five participants per year. Twenty laboratories from 17 countries have participated over the years, including Austria, Belgium, Canada, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Latvia, Norway, Russia, Sweden, Switzerland and UK. This year, four laboratories participated in the intercalibration with one taxonomist participating from each laboratory, including two new taxonomists.

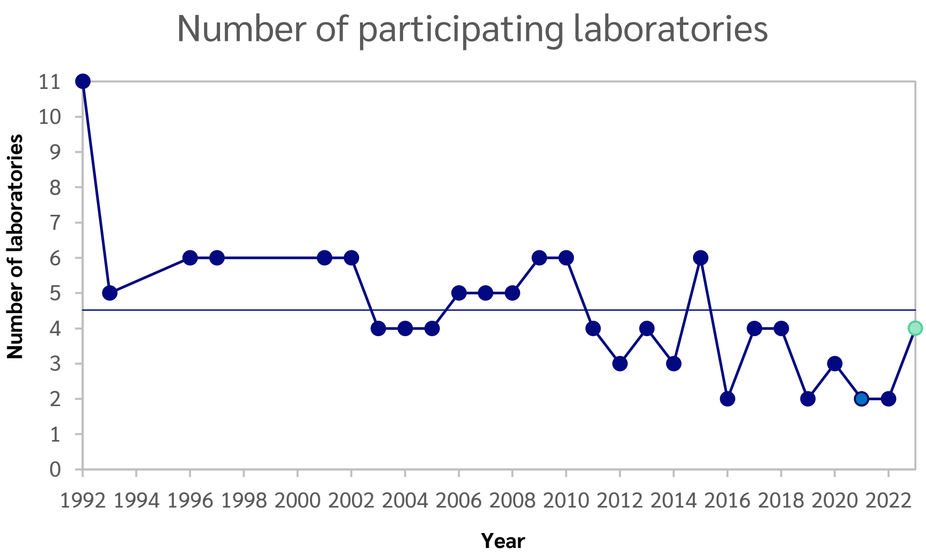


Figure 8. *The number of participating laboratories in the ICP Waters invertebrate intercalibration since the first intercalibration in 1992. The number of participants in 2023 is shown in green.*

The intercalibration laboratory protocol is unchanged since 1992, while the quality assurance index (Qi) has been used since it was introduced in 2005 (Raddum, 2005). After back calculating the Qi for the period prior to 2005, the Qi now is available from 1992 and up to the present (**Figure 9**). Trends in the Qi-score show that the mean has remained above 80%, suggesting good taxonomic work and skilled taxonomists in the laboratories affiliated to ICP Waters. When the Qi is broken into individual invertebrate groups, the laboratories, on average over the years, perform best for caddisflies and worst for stoneflies (**Figure 10**). The last two intercalibration results have shown a higher overall score in the plecopteran group than for the trichopterans, deviating from the previous years.

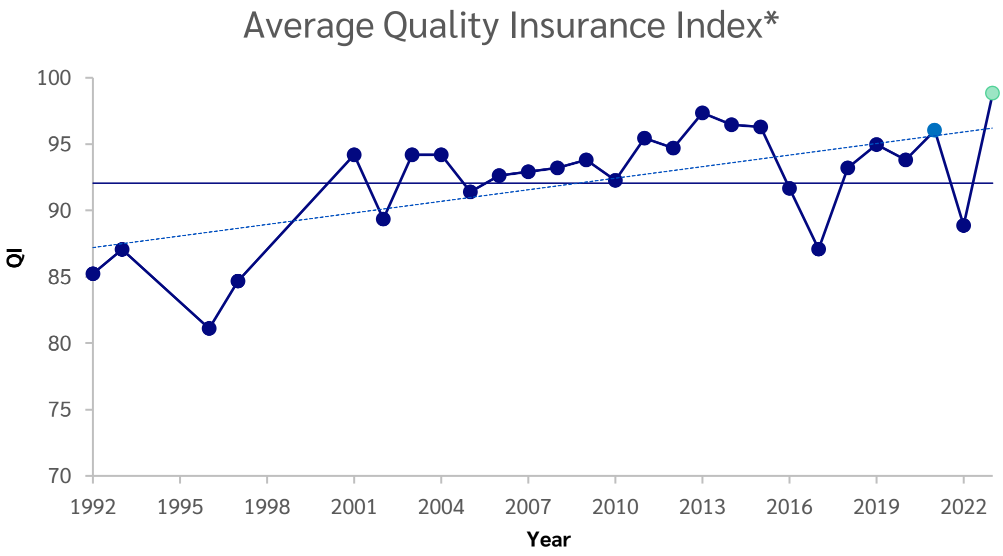
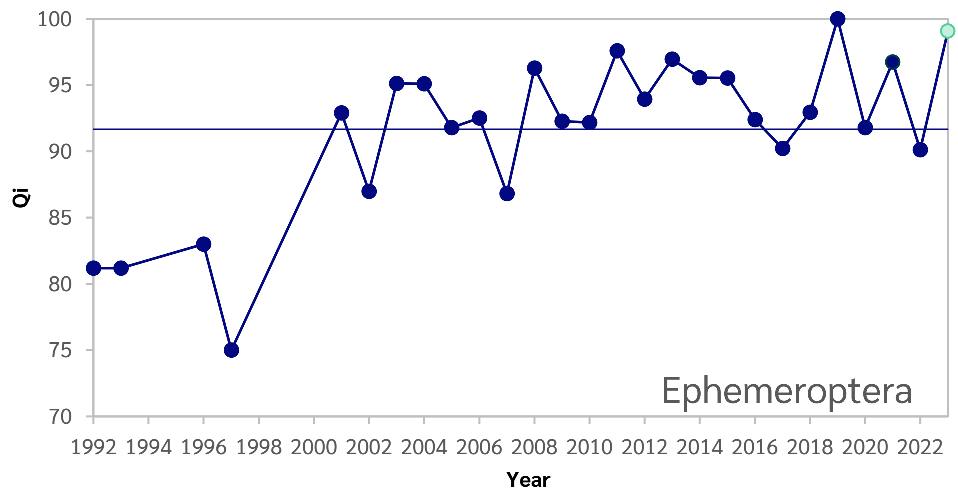


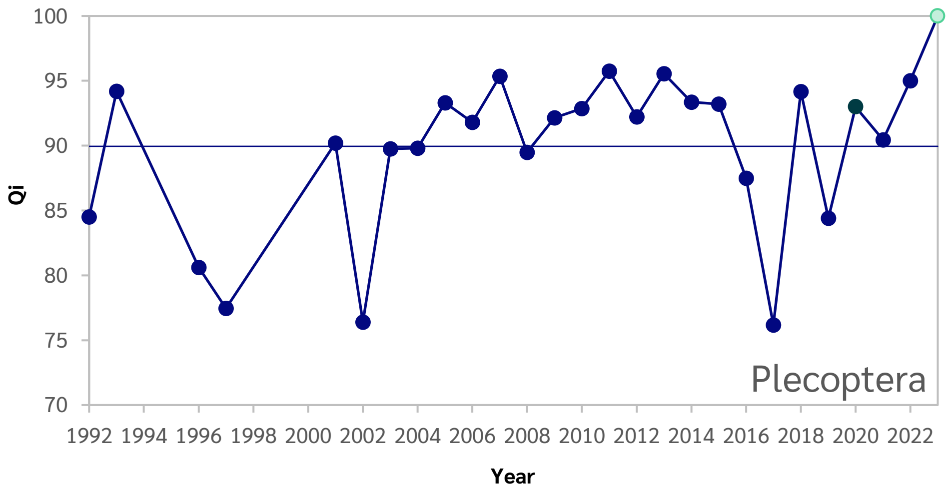
Figure 9. *The mean quality assurance index for the invertebrate intercalibration through time. Horizontal line represents the mean quality assurance index (QI) for the last 31 years. Results from 2023 are shown in green.*

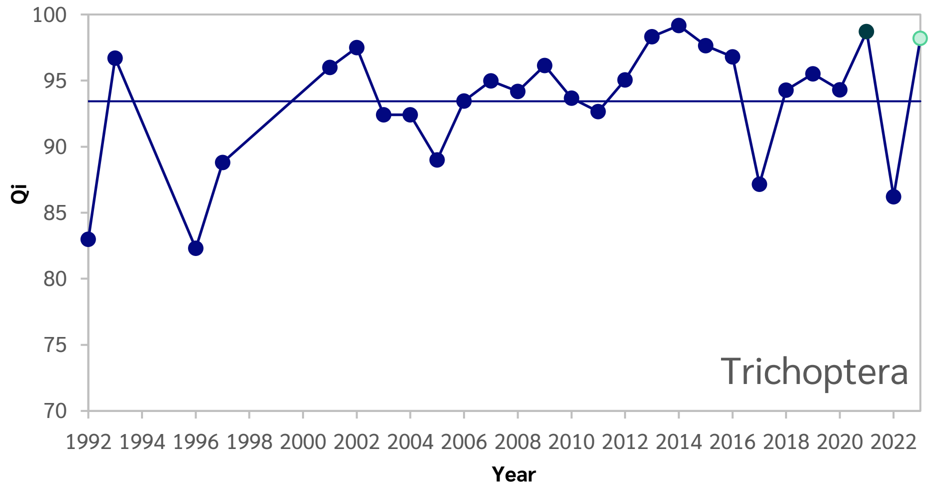
One of the aims of the intercalibration is to improve the taxonomic skill of the participating laboratories. The mean Qi has increased since the intercalibration started, suggesting that the skills have improved (**Figure 9**). Still, at least four issues influence the Qi:

1) The Qi varies according to the skills of the participants. A consequence is that the Qi often decreases when new laboratories participate or if a skilled taxonomist retires. As an example, the expert on the miscellaneous group retired from Laboratory 2 in 2018, which resulted in a low Qi.
2) The Qi varies according to the difficulty of the test, which mostly depends on specimen size and the rarity of the species. For example, more species in the miscellaneous group were included in the intercalibration around 2005 since new acidification indices demanded a higher taxonomic resolution for this group. Hence, the Qi subsequently dropped for some years before it gradually increased (**Figure 10**). The increase likely reflects improved taxonomic skill.
3) There is inevitably some chance involved. For example, samples have occasionally dried out, a taxonomist may have overlooked a specimen or forgotten to make comments on a damaged specimen.
4) Some years, the participants send too few specimens from their home region to the intercalibration organizer. This may influence the results since the organizer then needs to include specimen from other regions to the test of that specific participant. It is therefore important that the participants send an abundance of specimens to the organizer.
5) The mean Qi is calculated as the average of the scores from each taxonomic group. The Qi-score for each group is calculated from the percentage of errors made in the group. This means that a taxonomic error in a group with few individuals will have a larger negative impact on the Qi-score than an error in a group with many individuals.

The mean Qi has decreased during 2012-2017, more steeply between 2015 and 2017, to increase again towards the present. According to the taxonomists, the difficulty increased during 2015-2017, and especially for stoneflies. In addition, it seems some other above-mentioned factors apply; there was a new participant, one key taxonomist retired, one sample dried out and one laboratory sent too few specimens from their home region. Hopefully, the number of such events will decline during forthcoming intercalibrations as continuous communication with participants will improve the overall experience for both participants and organizers. The mean Qi for 2023 increased from 2022 and the average score from this year’s test was the highest achieved since the intercalibration started. Two taxonomists participated for the first time, showing excellent skill by maintaining a high score and contributing to the high overall score. In past years, trends indicate increased taxonomic skill by the participants, underlined by the high score of the 2023 intercalibration and its four laboratories. Highly skilled macroinvertebrate taxonomists play a pivotal role in biological work due to their expertise in identifying and classifying aquatic invertebrates. This year’s participants have displayed high quality taxonomic work in all categories. The precision of taxonomic assessments by skilled experts ensures reliable and consistent data, essential for informed decision-making in environmental conservation and management efforts.







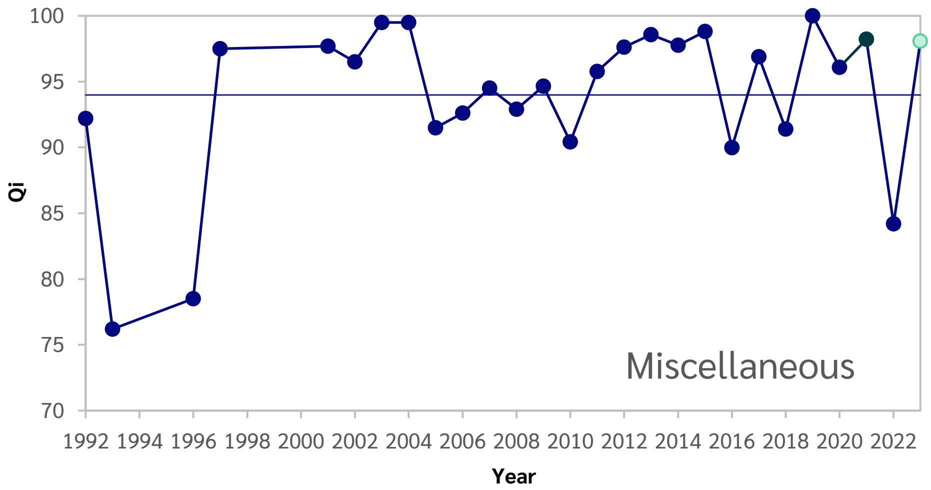


Figure 10. *The mean quality assurance index (Qi) of the intercalibrations through time for mayflies (Ephemeroptera), stoneflies (Plecoptera), caddisflies (Trichoptera) and miscellaneous groups of invertebrates. The horizontal line represents the overall mean Qi for each invertebrate group. The green marker indicates results from 2023. Qi above 80 is regarded as good taxonomical work.*

6 References

Armitage, P. D., D Moss, J. F. Wright and M. T. Furse, 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: 333–347.

Fauna Europaea Web Service (2013) Fauna Europaea version 2.6.2, Available online at http://www.faunaeur.org

Fjellheim, A. and G. G. Raddum, 1990. Acid precipitation: biological monitoring of streams and lakes. The Science of the Total Environment, 96, 57-66.

Halvorsen, G. A., E. Heegaard and G.G. Raddum, 2002. Tracing recovery from acidification – a multivariate approach. NIVA- Report SNO 4208/2000, ICP Waters Report 69/2002, 34 pp.

Halvorsen, G.A., Heegaard, E., Fjellheim, A. & Raddum, G.G. 2003. Tracing recovery from acidification in the Western Norwegian Nausta watershed. Ambio, 32(3): 234-239.

Larsen, J., H.J.B. Birks, G.G. Raddum & A. Fjellheim 1996. Quantitative relationships of invertebrates to pH in Norwegian river systems. Hydrobiologia 328: 57-74.

Raddum, G. G. 1999. Large scale monitoring of invertebrates: Aims, possibilities and acidification indexes. In Raddum, G. G., Rosseland, B. O. & Bowman, J. (eds.) Workshop on biological assessment and monitoring; evaluation of models. ICP-Waters Report 50/99, pp.7-16, NIVA, Oslo.

Raddum, G. G. 2005. Biological intercalibration: Invertebrates 0905. NIVA-report SNO 5067 2005, ICP Waters report 81/2005.

Raddum, G. G., A. Fjellheim and T. Hesthagen, 1988. Monitoring of acidification through the use of aquatic organisms. Verh. Int. Verein. Limnol. 23: 2291-2297.

Schartau, A.K., Moe, J., Sandin, L., McFarland, B. and Raddum, G. G. 2008. Macroinvertebrate indicators of lake acidification: analysis of monitoring data from UK, Norway and Sweden. Aquatic Ecology, 42: 293–305.

Skjelkvåle, B. L.; Andersen, T.; Halvorsen, G. A.; Raddum, G. G.; Heegaard, E.; Stoddard, J. and Wright, R. F. 2000. The 12-year report: Acidification of Surface Water in Europe and North America; Trends, biological recovery and heavy metals. ICP Waters report, nr. 52/2000. Oslo: Norwegian Institute for Water Research; 2000. 115 pp.

Velle, G., Telford, R. J., Curtis, C., Erikson, L., Fjellheim, A., Frolova, M., Fölster, J., Grudule, N., Halvorsen, G. A., Hildrew, A., Hoffmann, A., Inderiksone, I., Kamasová, L., Kopàček, Orton, S., Krám, P., Monteith, D. T., Senoo, T., Shilland, E. M., Stuchlik, E., Wiklund, M-L., deWit, H. and Skjelkvaale, B. L. 2013. Biodiversity in freshwaters: temporal trends and response to water chemistry. ICP Waters Report 6580-2013, Norwegian Institute for Water Research, Oslo.

Velle, G., Mahlum, S., Monteith, D., de Wit, H., Arle, J., Eriksson, L., Fjellheim, A., Frolova, M., Fölster, J., Grudule, N., Halvorsen, G.A., Hildrew, A., Hruska, J., Indriksone, I., Kamasová, L., Kopacek, J., Krám, P., Orton, S., Senoo, T., Shilland, E.M., Stuchlik, E., Telford, R.J., Wiklund, Unfermanova, L., Wiklund & Wright, R. 2016. Biodiversity of macro-invertebrates in acid-sensitive waters: trends and relations to water chemistry and climate. ICP Waters report 127/2016. 33 pp.

Velle, G., Bækkelie. K.A., Arnekleiv, J.V., Bongard, T., Bremnes, T., Hall, J., Halvorsen, G.A., Dahl-Hansen, I., Johansen, A., Kjærstad, G., Landås, T., Saltveit, S.J., Stabell, T. 2018. Kvalitetssikring av bunndyrundersøkelser i Norge. LFI report 315. 31 pp.

Velle, G., Birkeland, I.B., Landås, T., Johannessen, A. 2020. Ringtest for artsbestemmelser av bunndyr i ferskvann 2019. NORCE LFI rapport nr. 392, Miljødirektoratet rapport M-1824|2020. 36s.

Appendix A. Responsible laboratories

Each participating laboratory is identified by a number, which is identical with laboratory numbers in the report and Appendix B. Laboratories participating in the intercalibration of invertebrates in 2023 are:

1. EKOLOGIGRUPPEN AB Stora Sodergatan 8c LUND, 222 23, **Sweden**. Responsible taxonomist: Cecilia Holmström

2. Swedish University of Agricultural Sciences, Dept. of Environmental Assessment, P.O. Box 7050, S-75007 Uppsala, **Sweden**. Responsible taxonomist: Dr. Magda-Lena Wiklund.

3. EKUK - Estonian Environmental Research Centre Department of Tartu Vaksali 17a, 50410 Tartu, **Estonia**. Responsible taxonomist: Lilian Metsavas

4. Latvian Environment, Geology and Meteorology Centre Maskavas street 165, Latvia. Responsible taxonomist: Dāvis Ozoliņš

Appendix B. Species lists

Table 1. Identified species/genus in sample 1 and 2 by Laboratory 1

	Sample 1		Sample 2	
Groups	Delivered	Identified	Delivered	Identified
Ephemeroptera:				
Ephemera danica			1	1
Caenis rivulorum			2	2
Heptagenia sulphurea			1	1
Baetis buceratus			1	1
Baetis digitatus	1	1		
Baetis muticus	1	1		
Baetis niger	1	1		
Baetis rhodani	1	1		
Plecoptera:				
Brachyptera risi			2	2
Protonemoura meyeri			2	2
Amphinemura borealis			1	1
Amphinemura sulcicollis	1	1	1	1
Nemoura cinerea	1	1		
Leuctra nigra	1	1		
Isoperla grammatica	1	1		

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Trichoptera:				
Rhyacophila nubila			1	1
Chimarra marginata			1	1
Cheumatopsyche lepida	1	1		
Hydropsyche contubernalis			1	1
Hydropsyche saxonica			1	1
Hydropsyche siltalai			1	1
Agapetus ochripes	1	1		
Silo pallipes	1	1		
Sericostoma personatum	1	1		
Setodes argentipunctellus	1	1		
Miscellaneous:				
Turbellaria obest				
Polycelis sp.	1	1	1	1
Oligochaeta övriga				
Eiseniella tetraedra	1	1		
Hirudinea				
Helobdella stagnalis	1	1		
Erpobdella octoculata	1	1	1	1
Pisidium sp.			2	2
Sphaerium sp.	1	1		
Physa fontinalis	1	1		
Ancylus fluviatilis	1	1		
Acroloxus lacustris			1	1

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Miscellaneous:				
Potamopyrgus antipodarum			2	2
Asellus aquaticus			2	2
Onychogomphus forcipatus	1	1	1	1
Cordulegaster boltonii	1	1		
Aphelocheirus aestivalis	1	1		
Orectochilus villosus	1	1		
Hydraena gracilis	1	1		
Elmis aenea			1	1
Limnius volckmari			2	2
Normandia nitens			1	1
Oulimnius tuberculatus	1	1		
Stenelmis canaliculata	1	1	2	2
Sialis lutaria	1	1		
Eloeophila sp.	1	1		
Empididae	1	1		
Ibisia marginata	1	1		

Table 2. Identified species/genus in sample 1 and 2 by Laboratory 2

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Ephemeroptera:				
Baetis niger	1	1		
Ephemera vulgata	1	1		
Ephemerella aurivilli	1	1		
Heptagenia fuscogrisea	1	1		
Heptagenia sulphurea	1	1		
Baetis digitatus			2	2
Baetis muticus			2	2
Caenis horaria			2	2
Caenis rivulorum			1	1
Centroptilum luteolum			1	1
Plecoptera:				
Capnopsis schilleri	1	1		
Diura nanseni	1	1		
Nemoura flexuosa	1	1		
Siphonoperla burmeisteri	1	1	1	1
Taeniopteryx nebulosa	1	1		
Brachyptera risi			1	1
Leuctra nigra			1	1
Perlodes dispar			1	1

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Trichoptera:				
Athripsodes aterrimus	1	1		
Cyrnus trimaculatus	1	1		
Lepidostoma hirtum	1	1		
Micrasema setiferum	1	1		
Molanna angustata	1	1		
Polycentropus flavomaculatus	1	1		
Sericostoma personatum	1	1		
Silo pallipes	1	1		
Cheumatopsyche lepida			1	1
Ecnomus tenellus			1	1
Hydropsyche pellucidula			1	1
Micrasema gelidum			1	1
Molannodes tinctus			1	1
Mystacides azurea			1	1
Oecetis testacea			1	1
Tinodes waeneri			1	1
Miscellaneous:				
Acroloxus fluviatilis	1	1		
Bathyomphalus contortus	1	1		
Cordulegaster boltonii	1	1		
Limnius volckmari	1	1		
Limnophora sp.	1	1		
Elmis aenea	2	2		

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Miscellaneous:				
Asellus aquaticus			2	2
Dryops sp.			1	1
Gammarus pulex			2	2
Pyrrhosoma nymphula			1	1
Stenelmis canaliculata			2	2

Table 3. Identified species/genus in sample 1 and 2 by Laboratory 3

Groups	Sample 1 Delivered	Sample 1 Identified	Sample 2 Delivered	Sample 2 Identified
Ephemeroptera:				
Antropus fragilis	1	1		
Arthroplea congener	2	2		
Baetis buceratus	2	2		
Baetis fuscatus	1	1		
Baetis liebenauae	1	0		
Centroptilum luteolum	2	2		
Ephemerella mucronata	2	2	2	2
Serratella ignita	1	1		
Ephemera lineata			1	1
Electrogena affinis			1	1
Heptagenia dalecarlica			2	2
Habrophlebia fusca			1	1
Habrophlebia lauta			1	1
Potamanthus luteus			1	1
Siphlonorus aestivalis			2	2
Caenis horaria			2	2
Cleon dipterum			2	2
Plecoptera:				
Isoptena serricornis	2	2		
Siphonoperla burmeisteri	1	1		
Leuctra fusca	1	1		

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Plecoptera:				
Leuctra hippopus	2	2		
Amphinemura borealis	3	3		
Nemoura cinerea	1	1		
Nemoura flexuosa			1	1
Protonemura meyeri			2	2
Isoperla difformis			2	2
Isoperla grammatica			2	2
Perlodes dispar			1	1
Brachyptera risi			1	1
Trichoptera:				
Athripodes aterrimus			1	1
Ecnomus tenellus	2	2		
Chmiarra marginata	1	1		
Geora pilosa	1	1		
Silo nigricornis	1	1		
Silo Pallipes	1	1		
Cheimatopsyche lepida			2	2
Hydropsyche angustipennis			1	1
Hydropsyche pellucidula			2	2
Hydropsyche siltatai			1	1
Ceraclea excisa			1	1
Oecetis notata			1	1
Oecetis testacea			1	1

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Trichoptera:				
Ironoquia dubia			1	1
Limnephilus fuscicornis			1	1
Limnephilus lunatus			1	1
Potamophylax rotundipennis	1	1		
Dontocerum albicorne	2	2		
Phryganea bipunctata	1	1		
Cyrnus flavidus	2	2		
Trichostegia minor			1	1
Miscellaneous:				
Porhydrus lineatus	1	1		
Elmis aenea	1	1		
Elmis maugetii			1	1
Argyroneta aquatica			1	1
Paropynx stratiotata	1	1		
Cymatia bonsdorffii	1	1		
Sigara semistriata			1	1
Sigara striata			1	1
Mesovelia furcata			1	1
Notonecta glauca	1	1		
Nepa cinerea	1	1		
Sigara distincta	1	1		

Table 4. Identified species/genus in sample 1 and 2 by Laboratory 4

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Ephemeroptera:				
Alainites muticus	2	2		
Cloeon dipterum	2	2		
Heptagenia sulphurea	2	2		
Baetis digitatus	1	1		
Baetis liebenauae	1	1		
Baetis rhodani	1	1		
Siphlonurus lacustris	1	1		
Nigrobaetis niger	1	1		
Serratella ignita	1	1		
Ephemera danica	1	1	1	1
Baetis vernus			2	2
Brachycercus harrisella			2	2
Caenis lactea			1	2
Caenis luctuosa			2	1
Caenis robusta			1	1
Cenis horaria			2	2
Ephemera lineata			1	1
Ephemera vulgata			1	1

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Plecoptera:				
Amphinemura borealis	2	2		
Nemoura cinerea	1	1		
Brachyptera risi	1	1	1	1
Isoperla grammatica			2	2
Nemoura avicularis			2	2
Nemoura cinerea			1	1
Trichoptera:				
Athripsodes atterimus	2	2		
Beraeodes minutus	1	1		
Brachycentrus subnubilus	1	1		
Cyrnus flavidus	1	1	1	1
Hydropsyche angustipennis	1	1		
Ironoquia dubia	1	1	1	1
Lepidostoma hirtum	1	1		
Leptocerus tineiformis	1	1		
Brachycentrus subnubilus			1	1
Ceraclea annulicornis			1	0
Leptocerus tineiformis			1	1
Molanna angustata			2	2
Mystacides azurea			2	2

Groups	Sample 1 Delivered	Identified	Sample 2 Delivered	Identified
Trichoptera:				
Notidobia ciliaris			1	1
Oecetis testacea			1	1
Plectrocnemia conspersa			2	0
Polycentropus flavomaculatus			2	2
Potamophylax rotundipennis			1	1
Rhyacophila nubila			2	2
Miscellaneous:				
Bithynia leachii	1	1		
Bithynia tentaculata	1	1		
Valvata cristata	1	1		
Calopteryx virgo	1	1		
Cordulegaster boltonii	1	1		
Gomphus vulgatissimus	1	1		
Limnius volckmari	1	1		
Atherix ibis	1	1		
Gammarus lacustris	1	1	1	1
Gammarus pulex	1	0	1	1
Sialis sordida			1	1

Appendix C. Thematic reports from the ICP Waters programme

Since its establishment in 1985, the ICP Waters programme has prepared numerous assessments, reports and publications that address the effects of long-range transported air pollution, including thematic reports, chemical intercalibrations, biological intercalibrations, proceedings of Task Force meetings, and peer-reviewed articles.

Reports and publications are available at the ICP Waters website; http://www.icp-waters.no/

Thematic reports from the ICP Waters programme from 2000 up to present are listed below.

Velle, G.; Bodin, C.L.; Arle, J.; Austnes, K.; Boggero, A.; Bojkova, J.; Fornaroli, R.; Fölster, J.; Goedkoop, W.; Jones, I.; Juggins, S.; Lau, D.C.P.; Monteith, D.; Murphy, J.; Musazzi, S.; Shilland, E.; Steingruber, S.; Wiklund, M.-L.; de Wit, H. 2023. Responses of benthic invertebrates to chemical recovery from acidification. NIVA SNO 7881-2023. **ICP Waters report 153/2023**.

Austnes, K., Hjermann, D.Ø., Sample, J., Wright, R. F., Kaste, Ø., and de Wit, H. 2022. Nitrogen in surface waters: time trends and geographical patterns explained by deposition levels and catchment characteristics. NIVA SNO 7728-2022. **ICP Waters report 149/2022**.

Thrane, J.E., de Wit, H. and Austnes, K. 2021. Effects of nitrogen on nutrient-limitation in oligotrophic northern surface waters. NIVA report SNO 7680-2021. ICP Waters report 146/2021.

Garmo, Ø., Arle, J., Austnes, K. de Wit, H., Fölster, J., Houle, D., Hruška, J., Indriksone, I., Monteith, D., Rogora, M., Sample, J.E., Steingruber, S., Stoddard, J.L., Talkop, R., Trodd, W., Ulańczyk, R.P. and Vuorenmaa, J. 2020. Trends and patterns in surface water chemistry in Europe and North America between 1990 and 2016, with particular focus on changes in land use as a confounding factor for recovery. NIVA report SNO 7479-2020. **ICP Waters report 142/2020**.

Austnes, K. Aherne, J., Arle, J., Čičendajeva, M., Couture, S., Fölster, J., Garmo, Ø., Hruška, J., Monteith, D., Posch, M., Rogora, M., Sample, J., Skjelkvåle, B.L., Steingruber, S., Stoddard, J.L., Ulańczyk, R., van Dam, H., Velasco, M.T., Vuorenmaa, J., Wright, R.F., de Wit, H. 2018. Regional assessment of the current extent of acidification of surface waters in Europe and North America. NIVA report SNO 7268-2018. **ICP Waters report 135/2018**

Braaten, H.F.V., Åkerblom, S., de Wit, H.A., Skotte, G., Rask, M., Vuorenmaa, J., Kahilainen, K.K., Malinen, T., Rognerud, S., Lydersen, E., Amundsen, P.A., Kashulin, N., Kashulina, T., Terentyev, P., Christensen, G., Jackson-Blake, L., Lund, E. and Rosseland, B.O. 2017. Spatial and temporal trends of mercury in freshwater fish in Fennoscandia (1965-2015). NIVA report SNO 7179-2017. **ICP Waters report 132/2017**.

Velle, G., Mahlum, S., Monteith, D.T., de Wit, H., Arle, J., Eriksson, L., Fjellheim, A., Frolova, M., Fölster, J., Grudule, N., Halvorsen, G.A., Hildrew, A., Hruška, J., Indriksone, I., Kamasová, L., Kopáček, J., Krám, P., Orton, S., Senoo, T., Shilland, E.M., Stuchlík, E., Telford, R.J., Ungermanová, L., Wiklund, M.-L. and Wright, R.F. 2016. Biodiversity of macro-invertebrates in acid-sensitive waters: trends and relations to water chemistry and climate. NIVA report SNO 7077-2016. NIVA report SNO 7077-2016. **ICP Waters report 127/2016**.

De Wit, H., Hettelingh, J.P. and Harmens, H. 2015. Trends in ecosystem and health responses to long-range transported atmospheric pollutants. NIVA report SNO 6946-2015. **ICP Waters report 125/2015**.

De Wit, H. A., Garmo Ø. A. and Fjellheim A. 2015. Chemical and biological recovery in acid-sensitive waters: trends and prognosis. **ICP Waters Report 119/2014**.

Holen, S., R.F. Wright and Seifert, I. 2013. Effects of long-range transported air pollution (LTRAP) on freshwater ecosystem services. NIVA report SNO 6561-2013. **ICP Waters Report 115/2013**.

Velle, G., Telford, R.J., Curtis, C., Eriksson, L., Fjellheim, A., Frolova, M., Fölster J., Grudule N., Halvorsen G.A., Hildrew A., Hoffmann A., Indriksone I., Kamasová L., Kopáček J., Orton S., Krám P., Monteith D.T., Senoo T., Shilland E.M., Stuchlík E., Wiklund M.L., de Wit, H. and Skjelkvaale B.L. 2013. Biodiversity in freshwaters. Temporal trends and response to water chemistry. NIVA report SNO 6580-2013. **ICP Waters Report 114/2013**.

Wright, R.F., Helliwell, R., Hruska, J., Larssen, T., Rogora, M., Rzychoń, D., Skjelkvåle, B.L. and Worsztynowicz, A. 2011. Impacts of Air Pollution on Freshwater Acidification under Future Emission Reduction Scenarios; ICP Waters contribution to WGE report. NIVA report SNO 6243-2011. **ICP Waters report 108/2011**.

Skjelkvåle B.L. and de Wit, H. (eds.) 2011. Trends in precipitation chemistry, surface water chemistry and aquatic biota in acidified areas in Europe and North America from 1990 to 2008. NIVA report SNO 6218-2011. **ICP Waters report 106/2011**.

ICP Waters Programme Centre 2010. ICP Waters Programme manual. NIVA SNO 6074-2010. **ICP Waters report 105/2010**.

De Wit, H.A. and Lindholm M. 2010. Nutrient enrichment effects of atmospheric N deposition on biology in oligotrophic surface waters – a review. NIVA report SNO 6007 - 2010. **ICP Waters report 101/2010**.

Ranneklev, S.B., De Wit, H., Jenssen, M.T.S. and Skjelkvåle, B.L. 2009. An assessment of Hg in the freshwater aquatic environment related to long-range transported air pollution in Europe and North America. NIVA report SNO 5844-2009. **ICP Waters report 97/2009**.

Skjelkvåle, B.L., and De Wit, H. (eds.) 2008. ICP Waters 20 year with monitoring effects of long-range transboundary air pollution on surface waters in Europe and North-America. NIVA report SNO 5684-2008. **ICP Waters report 94/2008**.

Wright, R.F., Posch, M., Cosby, B. J., Forsius, M., and Skjelkvåle, B. L. 2007. Review of the Gothenburg Protocol: Chemical and biological responses in surface waters and soils. NIVA report SNO 5475-2007. **ICP Waters report 89/2007**.

Skjelkvåle, B.L., Forsius, M., Wright, R.F., de Wit, H., Raddum, G.G., and Sjøeng, A.S.M. 2006. Joint Workshop on Confounding Factors in Recovery from Acid Deposition in Surface Waters, 9-10 October 2006, Bergen, Norway; Summary and Abstracts. NIVA report SNO 5310-2006. **ICP Waters report 88/2006**.

De Wit, H. and Skjelkvåle, B.L. (eds) 2007. Trends in surface water chemistry and biota; The importance of confounding factors. NIVA report SNO 5385-2007. **ICP Waters report 87/2007**.

Wright, R.F., Cosby, B.J., Høgåsen, T., Larssen, T. and Posch, M. 2005. Critical Loads, Target Load Functions and Dynamic Modelling for Surface Waters and ICP Waters Sites. NIVA report SNO 5166-2005. **ICP Waters report 83/2006**.

Fjeld, E., Le Gall, A.-C. and Skjelkvåle, B.L. 2005. An assessment of POPs related to long-range air pollution in the aquatic environment. NIVA report SNO 5107-2005. **ICP Waters report 79/2005**.

Raddum, G.G, et al. 2004. Recovery from acidification of invertebrate fauna in ICP Water sites in Europe and North America. NIVA report SNO 4864-2004. **ICP Waters report 75/2004.**

Skjelkvåle, B.L. (ed) 2003. The 15-year report: Assessment and monitoring of surface waters in Europe and North America; acidification and recovery, dynamic modelling and heavy metals. NIVA report SNO 4716-2003. **ICP Waters report 73/2003.**

Wright, R.F. and Lie, M.C. 2002.Workshop on models for Biological Recovery from Acidification in a Changing Climate. 9-11 september 2002 in Grimstad, Norway. Workshop report. NIVA report 4589-2002.

Jenkins, A. Larssen, Th., Moldan, F., Posch, M. and Wrigth, R.F. 2002. Dynamic Modelling of Surface Waters: Impact of emission reduction - possibilities and limitations. NIVA report SNO 4598-2002. **ICP Waters report 70/2002**.

Halvorsen, G.A, Heergaard, E. and Raddum, G.G. 2002. Tracing recovery from acidification - a multivariate approach. NIVA report SNO 4564-2002. **ICP Waters report 69/2002**.

Wright, R.F. 2001. Note on: Effect of year-to-year variations in climate on trends in acidification. NIVA report SNO 4328-2001. **ICP Waters report 57/2001**.

Hovind, H. 2000. Trends in intercomparisons 8701-9812: pH, K_{25}, NO_3 + NO_2, Cl, SO_4, Ca, Mg, Na, K and aluminium - reactive and nonlabile, TOC, COD-Mn. NIVA report SNO 4281-2000, **ICP Waters Report 56/2000**.

Skjelkvåle, B.L., Olendrzynski, K., Stoddard, J., Traaen, T.S, Tarrason, L., Tørseth, K., Windjusveen, S. and Wright, R.F. 2001. Assessment of trends and leaching in Nitrogen at ICP Waters Sites (Europe And North America). NIVA report SNO 4383-2001. **ICP Waters report 54/2001.**

Skjelkvåle, B. L., Andersen, T., Halvorsen, G. A., Raddum, G.G., Heegaard, E., Stoddard, J. L., and Wright, R. F. 2000. The 12-year report; Acidification of Surface Water in Europe and North America; Trends, biological recovery and heavy metals. NIVA report SNO 4208/2000. **ICP Waters report 52/2000**.

The Norwegian Institute for Water Research

We are Norway's premier research institute in the fields of water and the environment. We are experts on ecosystems in both freshwater and marine environments, from mountains, lakes and rivers, to fjords, coasts and oceans. We develop science-based knowledge and solutions to challenges related to the interaction between water and climate, the environment, nature, people, resources and society.