# 2013 Freshwater Eionet Workshop - 19/20 September 2013, Copenhagen

## Session 2: Data quality and aggregation

## Document 2e: Quality and representativity SoE biology data: representativity of stations, reporting in correct scale (EQR vs. original metric scale), national classification systems by Jannicke Moe, ETC/ICM

### Background

SoE biological data from rivers and lakes is a new dataflow that was started in 2011 after two years of test reporting. The biological data include ecological quality ratio (EQR) and status class for the biological quality elements phytobenthos and macroinvertebrates in rivers and phytoplankton and macrophytes in lakes. In addition, selected metrics from lakes are included (e.g. proportion of cyanobacteria). The data reported are stored and published in Waterbase: [http://www.eea.europa.eu/data-and-maps/data/waterbase-rivers-9](http://www.eea.europa.eu/data-and-maps/data/waterbase-rivers-8)

<http://www.eea.europa.eu/data-and-maps/data/waterbase-lakes-8>, and include data from 25 countries. The added value of these data relative to the WFD reporting is that they can be linked to specific pressures and impacts (e.g. nutrients/eutrophication, organic pollution). After some more years of reporting these data can also be used to illustrate temporal trends and potentially show the ecological effect of measures to mitigate the specific pressures. Moreover, coupling to other types of data (e.g. resource efficiency, emissions) can be used to further explore pressure-status-impact-response relationships, including links to ecosystem services and biodiversity. The assessment so far includes visualisation in WISE maps (<http://www.eea.europa.eu/themes/water/interactive/biological-water-quality-in-rivers>) and a draft assessment, which will be distributed prior to the Eionet meeting.

### Some suggested data issues to be presented and discussed

* **Data quality:** Currently, one third of the countries report biological indices on a wrong scale (e.g. absolute scale for national metrics rather than EQR scale and vice versa for the additional lake parameters, e.g. chlorophyll a or proportion Cyanobacteria). Correct scale is necessary for further European assessments and for comparisons between countries. Some countries report only status class without EQR values. Some countries report EQR values but do not report their classification system correctly. This prevents normalisation of the EQR values by ETC and thereby the use of the data for more detailed assessments. While status class provides a coarse estimate of ecological status for the various biological quality elements, the normalised EQR values are needed for trend analysis. Due to these quality issues, a substantial proportion of the data reported cannot be used for detailed assessment.
* **Data quantity and representativity:** A low number of stations are reported by one third of the countries reporting biological data. The representativity of these stations is unclear in terms of geographical cover, water body types and status classes. Some results based on SoE data show quite large deviations in the distribution of status classes compared to the WFD data for the same biological quality element (see background document with example). This can be due to different years used for the SoE stations than for the WFD reporting, but can also suggest a problem with representativity of the SoE stations. In order to assess the representativity of the data it is important that the water body code as reported to the WFD is supplied in the stations table for the SoE stations.

### Questions to NRCs

1. **Data quality:** Does the feedback from ETC uploaded to the countries' CDR folders provide sufficient information on why data are excluded from WISE maps and assessments? How is the QA procedure in your country before data reporting? Is the most recent progress in the development and intercalibration of classification systems for national metrics taken into account? What information do the countries need to better clarify the correct scale of reporting?
2. **Data quantity/representativity:** Are the reported SoE stations representative in terms of geographical coverage, water body types and status classes? If not, how can this be improved?
3. **Additional parameters:** Is there added value of including more parameters in the WISE maps? (the additional parameters that are already reported, but so far not used in the maps, e.g. lake chlorophyll and proportion Cyanobacteria).
4. **Length of time series:** Do you have old data of the same national biological metrics that have so far not been reported (from existing SoE stations or other, still monitored stations)? To allow trend analysis of the biological data as soon as possible those old data would be highly appreciated.
5. **Data Dictionary:** Would the development of a tailor-made input tool make the complicated reporting easier for you? If yes, how should the tool function?