

**UPGRADE BLACK SEA SCENE**

**GA 226592, FP7, EC**



**GUIDELINES FOR QUALITY CONTROL OF BIOLOGICAL DATA**

**PHYTOPLANKTON**

**Snejana Moncheva**

**November 2010**

This document is prepared under UP –Grade Black Sea Scene Project, No 226592, EC, FP7 as an essential tool to assure that the online database infrastructure contains high quality, harmonised data sets to serve research and as a basis for environmental policy and strategy development, decision making and management of the Black Sea ecosystem.

The document has been discussed and adopted during the *ad hoc Black Sea* Expert phytoplankton group meeting held in Istanbul, 22-23 June 2010, organized by Black Sea Commission, under the UP-Grade BS Scene Project.

The contribution of all participants is greatly acknowledged (*List of participants*)

	<b>Page</b>
<b>1 INTRODUCTION</b>	<b>4</b>
<b>2 THE QUALITY SYSTEM FOR BIOLOGICAL DATA - QA/QC</b>	<b>4</b>
<b>3 QA/QC FOR FIELD SAMPLING - SAMPLE PRESERVATION AND STORAGE</b>	<b>6</b>
Equipment	
Sampling Protocol	
Sample Preservation	
Sub-Sampling – Validation of Homogenization	
<b>4 QUALITY ASSESSMENT FOR LABORATORY ANALYSIS</b>	<b>6</b>
Taxonomy	
Cell Counts	
Biovolume/Biomass Estimation	
Re-Analysis	
Repeatability and Reproducibility	
Uncertainty	
Control Charts for Biological Measurements	
Training and Inter-Laboratory Comparability Testing	
<b>5 QUALITY ASSURANCE OF DATA REPORTING</b>	<b>12</b>
Documentation	
Data management	
MetaData Reporting Form	
<b>6. DATA FLAGGING SYSTEM</b>	<b>14</b>
<b>7. CONCLUSIONS</b>	<b>15</b>
<b>8. REFERENCES</b>	<b>16</b>
<b>9. LIST OF PARTICIPANTS</b>	<b>17</b>

## 1 INTRODUCTION

QA/QC of biological data has gained a recognition as an essential part of international monitoring programmes, in response to the demand for strategic environmental evaluations. The EU Water Framework Directive and the Marine Strategy Framework Directive presently provide a strong stimulus for the harmonisation of monitoring methods used by member states, and the quality control of the resulting data implemented. Further more the “ecosystem approach” will, if successfully implemented, have profound implications for future assessment strategies and, at the very least, is likely to require closer integration of data sets from disciplines which have conventionally been gathered and reported separately, and therefore have little in common with regard to QA/QC approaches. Biodiversity conservation, and monitoring is central in the Bucharest Convention and the BSC activity towards Black Sea environment protection. Informed decisions for environmental sound management can be made only on the basis of reliable data, and therefore certain level of data quality should be achieved to assure accuracy and precision of all measurement systems. Standards such as the ISO 9000 series and ISO 17025 provide a general framework for quality assurance but so far criteria for determining the acceptability of data from surveys of biological communities to meet specified information needs at international level are still under development, and should be given high priority.

## 2. THE QUALITY SYSTEM FOR BIOLOGICAL DATA - QA/QC

By definition *Quality assurance* includes all those planned and systematic actions necessary to provide adequate confidence so a product will satisfy given requirements of quality. This includes AQC, audit, training, documentation of methods, calibration schedules, etc., while *Quality control* stands for the operational techniques and activities that are used to fulfil requirements for quality. This is done by periodic measurement of quality of the product.

*The objective of a quality assurance programme is to reduce analytical errors to required limits and to assure that the results have a high probability of being of acceptable quality.*

Having developed an analytical system suitable for producing analytical results of the required accuracy, it is of eminent importance to establish a continuous control over the system and to show that all causes of errors remain the same in routine analyses (i.e., that the results are meaningful). In other words, continuous quantitative experimental evidence must be provided in order to demonstrate that the stated performance characteristics of the method chosen remain constant. To help meet the needs of the Black Sea laboratories to improve the quality of analytical results, a widely accepted standard for biological sampling and sample treatment/analysis is required. Such a standard cannot give a single best protocol as the questions underlying monitoring programmes are too diverse in character.

*This guidance is intended, therefore, to provide a uniform framework to optimize and validate ‘house’ methods and protocols in a comparable and standardized way. In order to reduce the high variance in the results of biological analysis, **standardization of as many steps as possible of the analytical procedures is necessary.***

The bounds are given by

- requirements from legislation
- standards of analytical methods and requirements for internal quality control (IQC)

- the (at least) laboratory-specific precision and trueness of the analytical value, which had to be ensured
- the valuation of laboratory-intern known data of the same sample type

The basic principle is that the validity of biological and ecological assessments depends on the accuracy and precision of all activities involved in the collection and analysis of data. Major variables include the characteristics of the taxonomic groups, the number of observations or measurements, their statistical distribution, the accuracy of identification guides, measuring devices or other methods and the skill of the surveyors or analysts in using these. This standard describes procedures for ensuring the quality of biological and ecological assessments of aquatic environments.

*These procedures encompass:*

- study design,
- surveying and sampling,
- analysis and identification,
- computing,
- validation,
- data interpretation and reporting,
- and training of personnel.

## **PHYTOPLANKTON**

The structural characteristics of phytoplankton communities bear valuable information about the evolution of microalgal communities and the trajectories of shifts under multiple environmental factors, including anthropogenic impacts. Irrespective of the available manual for phytoplankton sampling and analysis in the Black Sea (Moncheva, Parr, 2005) based on agreed procedures among laboratories from the 6 Black Sea countries and the outcome of an inter-comparison exercise (Black Sea Recovery Project, 2005) the procedures are not fully followed, or labs are working according to their own routines. Details of phytoplankton analytical procedures are essential to compare data produced by different analysts either during a long-term monitoring programs in one area or between different areas in order to evaluate statistically significant long-term trends or spatial differences. Carbon biomass of planktonic organisms is a fundamental parameter in ecosystem models and biogeochemical carbon budgets. Temporal and spatial variability in total and export primary production can be quantified and predicted only if the carbon content of the major plankton organisms is known. Carbon is the principal structural component of both heterotrophic and phototrophic organisms and is the basis for community-wide as well as group-specific comparisons of biomass and bioenergetics. Estimates of carbon biomass of plankton organisms are usually made by converting microscopic size measurements to cell volumes, which are then expressed in carbon biomass using empirically or theoretically derived carbon to volume ratios (Menden-Deuer, Lessard, 2000). This assures comparability of data when regional data bases are composed and further used in various regional studies.

*For phytoplankton a detailed methodological provision of Equipment, Sampling, Sample Preservation and storage and Laboratory analysis is given in the “Manual for Black sea phytoplankton sampling and analysis” (Moncheva, Parr, 2005, updated 2010). Thus for some QC/QA reference is made to the corresponding chapters of the Manual.*

*For routines see the Manual, Chapter 2. SAMPLING and Chapter 3. SAMPLE PRESERVATION AND STORAGE*

Survey and sample records must include sufficient information to enable the location of the sample or survey site to be identified precisely by future workers e.g. geo-referencing for mapping. Appropriate information may include grid references, plus notes, sketch maps and photographs relating the sample site to permanent landmarks in the vicinity. Survey and sample records, must have unique identifiers linking them to the time and place of collection. Sufficient meta-data must be recorded to ensure traceability.

#### **4 QUALITY ASSESSMENT FOR LABORATORY ANALYSIS**

*(For routines see the Manual, Chapter 6. PHYTOPLANKTON ENUMERATION AND DENSITIES*

##### **Taxonomy**

Accurate identification is essential especially when establishing a data base for future reference. It is a prerequisite to follow the systematic nomenclature developed by WoRMs (<http://www.marinespecies.org>) for unification of taxonomic lists. Proper identification should be controlled by setting up a taxon lists depicting all taxa identified by the analysts of the laboratory, and a reference check list is essential. Any taxa that are uncommon in the waterbody/region etc. should be highlighted and expert confirmation sought.

Species checklist with reference list of biovolumes is one of the most important components of the Quality Control procedure for the phytoplankton data from Black Sea region. It will help to provide comparable and homogeneous data sets of phytoplankton for the Black Sea basin. The list of phytoplankton species was developed as dynamic web-site, available on-line: <http://phyto.bss.ibss.org.ua>

*For details see Manual Annex A. PHYTOPLANKTON CHECK LIST*

The skills of taxonomists is also critical. On a regular basis intra-laboratory comparison tests should be performed to avoid/minimize identification differences between analysts. In addition to nationally or internationally organised inter-laboratory studies, control with related laboratories (dealing with comparable phytoplankton communities) is highly recommended.

##### **Phytoplankton cell counts**

Counting of at least 400 units per sample is compulsory to reach 95% significance level. In the protocols the real number of cells counted should be reported and not the calculated numbers for the chamber or the sample. The most dominating species should exceed 50-100 individual cells counted to give statistically good results. However statistically valid targets for cell counts and the lower detection limits of counted cells are still a major subject of standardization (EN 15204).

## **Biovolume [ $\mu\text{m}^3$ ]/ biomass estimation [ $\text{mg}/\text{m}^3$ ]**

The most commonly used traditional biomass calculation method for microalgae is cell biovolume, which is calculated from measured linear dimensions under microscope. Automated or semiautomated methods are not yet fully developed, and conventional light-microscope-based biovolume estimates are still in use. Few attempts have been made to standardize the calculation of algal biovolume. As a result, different sets of equations were used by different researchers. Efforts have been made to create one reference list of biovolumes for Black Sea microalgae. Thus for each species in the checklist the appropriate suggested figure to calculate biovolume was attached. For detailed research of morphometric characteristics of the community the more precise figure is also suggested where possible.

*For details see Manual Annex A. PHYTOPLANKTON CHECK LIST*

An automated system for phytoplankton data processing and quality control under development by IBSS to be used by all Black Sea countries will provide solution for the future. The different approaches used for biomass calculation is a serious source of differences. In the meta data it is essential to include precise information about the method used.

### ***Re-analysis***

Ten (10%) percent of all phytoplankton samples should be re-analysed by another analyst, and the results compared for QA purposes

### ***Repeatability and reproducibility***

The determination of the performance characteristics repeatability and reproducibility should be carried out with natural samples from a range of relevant conditions. For proper interpretation it should be stated whether the sample was rich or poor in algae, detritus etc.. Quantitative results should be assessed on a couple of most numerous species or, less preferably, on the level of the most common taxonomic groups (diatoms, dinoflagellates and others). The measured error should be compared with the expected stochastic error based on Poisson statistics to get insight into the performance of the procedure and analyst.

## **Quantitative uncertainty**

An abundance or composition estimate cannot be properly interpreted without knowledge of its uncertainty. Uncertainty of the final result of a microscopic analysis encompasses the uncertainties of the whole measurement process (sampling, stability, subsampling, homogeneity, identification, and quantification). A clear distinction should be made between quantitative and qualitative uncertainty. Qualitative uncertainty refers to mis- and non-identification of taxa. This part of uncertainty can only be dealt with in a general statement based on inter- (if available) and intra-laboratory comparisons. As yet there is no way to combine qualitative- and quantitative uncertainty estimates in one meaningful overall uncertainty. This should be part of a sensitivity analysis.

When it has been demonstrated by validation that the procedure of homogenization, subsampling and sedimentation has been brought into a state of statistical control, a meaningful uncertainty statement can be developed. Measurement uncertainty can be thought of as the sum of the intra-laboratory reproducibility and the trueness. Trueness is difficult to assess as the true value is actually always unknown. When available, trueness may be derived

from the deviation of the mean score in inter-laboratory studies. Another approach can be obtained by using some other method as for instance SR-chamber or electronic particle counters as a reference method. Inter-laboratory reproducibility can be assessed by recounting a representative set of samples under different relevant conditions.

*Quantitative uncertainty is dependent on the abundance of the specific taxon in the sample. This means that in the same sample the uncertainty of the abundance estimate of one taxon might be smaller or greater than that from another taxon. In general the uncertainty will increase with a decrease in abundance and the volume of the counting camera.*

$$\begin{aligned} X^2 = & \frac{(x_{11} - n_1 \hat{p}_1)^2}{n_1 \hat{p}_1} + \frac{(x_{21} - n_1 \hat{p}_2)^2}{n_1 \hat{p}_2} + \dots + \frac{(x_{k1} - n_1 \hat{p}_k)^2}{n_1 \hat{p}_k} \\ & + \frac{(x_{12} - n_2 \hat{p}_1)^2}{n_2 \hat{p}_1} + \frac{(x_{22} - n_2 \hat{p}_2)^2}{n_2 \hat{p}_2} + \dots + \frac{(x_{k2} - n_2 \hat{p}_k)^2}{n_2 \hat{p}_k} \quad (\text{df} = (N - 1)(k - 1)) \\ & + \frac{(x_{1N} - n_N \hat{p}_1)^2}{n_N \hat{p}_1} + \frac{(x_{2N} - n_N \hat{p}_2)^2}{n_N \hat{p}_2} + \dots + \frac{(x_{kN} - n_N \hat{p}_k)^2}{n_N \hat{p}_k} \end{aligned}$$

in which  $\hat{p}_i = x'_i / n'$  and

$$\sum x_i^2 / n_i = (x_{11}^2 / n_1) + (x_{12}^2 / n_2) + \dots + (x_{1N}^2 / n_N)$$

As a general rule it holds that if random samples of size n are drawn from a non-normal population the distribution of the mean will tend towards normal as n increases in size. This so-called central limit theorem can be applied to Poisson series when  $n > 30$  in which n is the number of grids and the mean number of a taxon or group per grid  $\bar{x}$ . In this case estimation of confidence limits for the estimate of the mean is straightforward according to:

$$\bar{x} \pm t_{0,05(2),v} \sqrt{\frac{\bar{x}}{n}}$$

in which n is the number of grids/fields counted per chamber,  $\bar{x}$  the mean number of a particular alga or group per grid/field and  $v = n - 1$ .

The confidence limits represent only the uncertainty in the estimation of the mean for a particular chamber. The reproducibility error should be added by counting different chambers etc. and then calculating the overall variance. Then the confidence limits can be calculated according to

$$\bar{x} \pm t_{0,05(2),v} \sqrt{\frac{S_p^2}{n + m}}$$

in which  $v = (n-1)+(m-1)$   $n$  is number of grids/fields counted per chamber,  $m$  is the number of chambers and  $s_p^2$  the overall variance which is the sum of  $\bar{x}$  and the reproducibility variance.  $s_R^2$ .

This variance should encompass, if relevant, different analysts, days, chambers etc. When  $n < 30$  the simplest way is to cluster those taxa which do not fulfil the constraint of  $n > 30$  until  $n > 30$  and then calculate the measurement uncertainty for this clustered group as a whole. Otherwise a direct method for the calculations of the confidence limits for the Poisson parameter can be applied.

The best approach to calculate **confidence limits** for the total count irrespective of the number of grids or objects counted is:

$$\text{Lower } 1-\alpha \text{ confidence limit } L_1 = \frac{\chi^2_{(1-\alpha/2),v}}{2}$$

in which  $v = 2X$

$$\text{Upper } 1-\alpha \text{ confidence limit } L_2 = \frac{\chi^2_{(\alpha/2),v}}{2}$$

in which  $v = 2(X+1)$

and  $X$  = number of objects counted.

This approach for calculating the confidence limits for the Poisson parameter results in an asymmetrical confidence interval. When a significant part of a chamber has been screened (for instance for large diatoms or desmids) the Poisson series is still applicable. The recommended procedure to carry out is to calculate the confidence limits according to the above mentioned method for calculating confidence limits for the Poisson parameter. To estimate the total number for the whole chamber the count and its confidence limits are multiplied by the ratio between sample and subsample size.

**NOTE** When dealing with colonies, it should be realized that the measurement uncertainty as calculated above, refers to the number of colonies and not to the number of cells. The number of cells in at least 30 replicate colonies should then be counted in order to estimate the mean.

*As for phytoplankton the absolute statistical limits are difficult to assess, particularly when no standards or other reference methods exist, in these cases the mean value obtained in interlaboratory studies organised among proficient laboratories can be adopted as practical limits.*

Bellow are given different option for **Total biomass target**

- The total value should be +/-20% of the value obtained from re-analysis of the sample.

- Bray-Curtis comparison: Comparison of the two untransformed data sets, arising from the work of the participating laboratory and from independent re-analysis, should result in a Bray-Curtis similarity index of  $\geq 90\%$ . For self-performance Bray-Curtis Similarity Index (BCSI:) scale is as follows:

100% BCSI:	Excellent
95-<100% BCSI:	Good
90-95% BCSI:	Acceptable
85-90% BCSI:	Poor – remedial action suggested
<85% BCSI:	Fail – remedial action required

- Phytoplankton analysis performed in a strict and standardized way have shown to generate coefficient of variations (CV) for biomass concentrations of about 27% for autotrophic microplankton and 18% for autotrophic nanoplankton (Andersson and Rudehäll 1993).

### ***Control Charts for Biological Measurements***

For quality control while measuring biological variables, the Shewhart charts where the criteria for evaluation of testing results is based on statistically calculated values, are used. The main advantage of Shewhart charts is the possibility to monitor the testing process; nevertheless, there is still the disadvantage that an accepted statistical deviation may be greater than the maximum deviation set by the method.

The control chart for duplicate samples can be constructed as follows: run one duplicate sample within every batch of samples. For phytoplankton, run every tenth sample or at least one sample per batch as duplicate, counting two subsamples from the same sample (approximately 10% of all samples). Calculate standard deviation  $S_R$  from the following equation

$$S_R = \pm \sqrt{\frac{\sum(|\Delta x_i|/2)^2}{n-1}}, \text{ where}$$

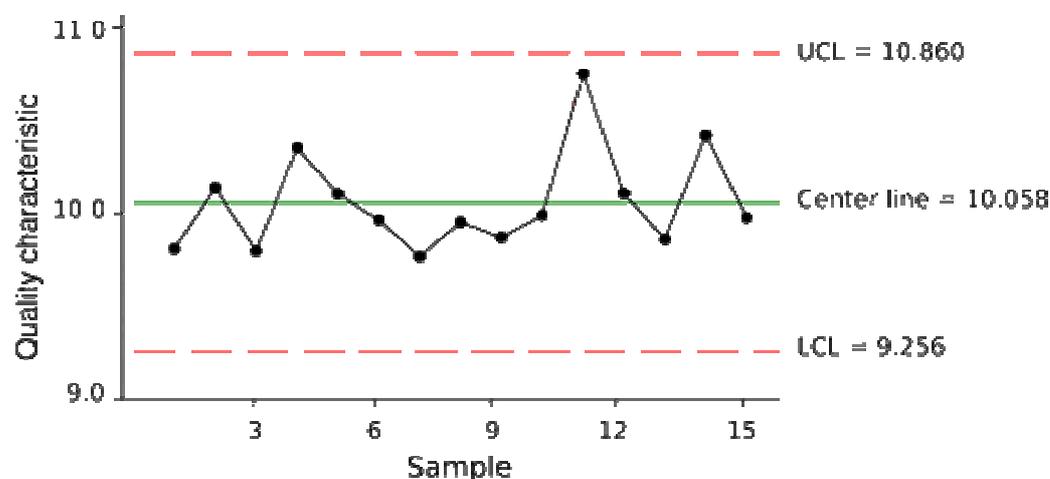
$n$  is the number of pairs of duplicate samples.

( $S_R$  can be calculated on the basis of testing results of duplicate samples taking into account at least ten analytical series.)

For QC plot the difference between testing results  $[\Delta x_i]/2$  versus time. Plot the standard deviation  $S_R$  vs. time. The construction of control charts can be done using any statistical software. The *warning limit of the analysis precision is two standard deviations*.

The quality control chart (Fig below) is intended to identify changes in random or systematic error. The following criteria for out-of-control situations are recommended for use with Shewhart charts:

- 1 control value being outside the action limits (upper) UA and [lower] LA; or
- 2 consecutive values outside warning limits UW and LW; or
- 7 consecutive control values with rising tendency; or
- 7 consecutive control values with falling tendency; or
- 10 out of 11 consecutive control values being on one side of the central line.



Shewhart charts example

(<http://en.wikipedia.org/wiki/File:ControlChart.svg>)

### *Stochastic errors*

The result of a measurement or assessment deviates from the true value because of the existence of a number of systematic and random errors. Particularly in biology important sources of random error are those introduced by the sampling and sub-sampling of biological items. These errors are stochastic by nature and should be considered separately. As one cannot do better than theoretical probability distributions predicts, the calculation of these errors is an important tool in the design of biological and ecological studies.

### *Elimination of systematic errors*

To check for systematic errors, several different trueness control samples are analyzed. To detect errors depending on the reagents or the method, control samples should be used whose concentrations cover the entire measuring range. As a minimum, a **trueness control** sample in the lower and one in the upper part of the working range should be

used. In the event of a systematic error with results predominantly being higher or lower than the actual values, a step by step examination should be performed to find the reason for this bias. Exchanging experimental parameters, such as reagents, apparatus or staff, might help to identify quickly this type.

#### *Improving precision*

The precision can also be improved by a step-by-step approach to find the causes of random error. The total precision of an analytical method can be improved by examining its individual procedural steps to find the one which contributes most to the total error.

#### **Plausibility control.**

There could be errors which may not be detected by a statistical approach to quality control. In most cases, this concerns errors influencing individual analyses in a batch. This type of error can only be revealed by means of *plausibility controls*:

*checks on the observed value in relation to expectations based on previous knowledge. A successful approach to plausibility control requires that appropriate background information is available. A possible approach is to construct a full year data matrix for several years in order to assist setting limits/ranges of possible variability (min-max values).*

#### ***Training and Inter-Laboratory Comparability Testing***

While the use of a validated analytical method and routine quality control (see above) will ensure accurate results within a laboratory, participation in an external quality assessment or proficiency testing scheme provides an independent and continuous means of detecting undiscovered sources of errors and demonstrate that the analytical quality control of the laboratory is effective. Important to improve and secure personal professional skills is participation and performance in Sampling and sample-handling Workshops, Taxonomic Training Workshops, Intercalibration exercises, Ring-tests.

## **5. QUALITY ASSURANCE OF DATA REPORTING**

### **Documentation**

All biological data produced by a laboratory should be completely documented (“metainformation”) and should be traceable back to its origin. The necessary documentation should contain a description of sampling equipment and procedures, reference to SOPs for the sampling, sample handling and analytical procedures involved, and the names of persons responsible for Quality Control. In general, one signed protocol should accompany a sample through all steps of processing.

**Problem:** *For biological data ODV format is not applicable. So far for the Black Sea the format developed by EUROOCEAN has been recommended and used (SESAME Project). There is an urgent need to develop adequate biological data base formats (stand alone document)*

## **Data management**

For the adequate management of the data obtained (especially when different laboratories are involved), an information management system is essential. The database should allow the storage/management of the full set of information relating to the data (including QA procedures, and summaries of analytical methods). A proper reporting format or data entry system should allow the submission of the required information in order to describe fully, and if necessary to trace back, the data/samples.

*Data checks performed by the (national) data manager should only be carried out on a data set that has already been subject to quality control procedures by the reporting institution.*

Therefore, information on QA/AQC procedures and outcomes has to accompany the data or, better, has to be regarded as part of the data submission (see below).

A central data management system should guarantee safe archiving (regular back-ups, computer virus checks, multiple storage, etc.) and access to the data.

*Check routines performed by the data management system should look for:*

- format compliance;
- completeness of data/information;
- compliance with the programme and guidelines;
- deviations from previous sampling/processing/analysis procedures;
- Cell volumes and dimensions
- plausibility (involving screening for outliers, e.g., arising from errors in position-fixing, or
- improbably high/low data values);
- conformity with agreed taxonomic nomenclature (parallel considerations include correct application of international coding systems, taxonomic updates, and synonyms);
- species occurrences additional to those in standard lists which may include non-native species.

“Quick-look” visualization of the data/information (e.g., in the form of track plots or charts) should be provided by the data centre, as well as meta-information relating to the submission of the data, including its state of validation

## **METADATA REPORTING FORM**

DATASET-NAME:

\*PROJECT:

\*PLATFORM/SHIP :

\*STATION COORDINATES:

\*GEOGRAPHIC-COVERAGE:

\*DATE & HOUR :

\*BOTTOM DEPTH [m]:

SAMPLING DEPTH [m]:

\*OBSERVED-PARAMETERS: (wind, currents, etc.)

\*MEASURED/DETERMINED-PARAMETERS: ( $T^{\circ}$ , salinity, in situ fluorescence, Secchi depth, pH,  $O_2$ ,  $H_2S$  etc.)

\*COMPUTED/CONVERTED-PARAMETERS: (inorganic nutrients, DOC, POC, PON, POM, chlorophyll a etc.)

## **SAMPLING INSTRUMENTS AND METHODS:**

Sampling instrument/equipment:

Sample volume :  
Sample preservation and handling:  
Responsible for data collection:

**DETAILED TAXONOMIC-IDENTIFICATION and ANALYSIS:**

Method of sample concentration:  
Volume of the sub-sample analyzed:  
Type of microscope used :  
Magnification used:  
Type of the counting chamber:  
Area of the chamber analyzed:  
Number of cells/ units counts:  
Biovolume determination method:  
C biomass estimation method:  
Identification system/books:  
Image library:  
Storage media:

**\*QUALITY CHECK Procedure (available basic documents):**

**Data entry (double entry):**

**\*TAXONOMISTS IN CHARGE OF THE DETERMINATION:**

(take part in sampling and sample-handling Workshops, Taxonomic training Workshops, Intercalibrations , in ring-tests).

## **6. DATA FLAGGING SYSTEM**

### **Data values, the problem of outliers**

Generally there are three key questions associated to the data values:

- Have the appropriate standards been met by the bulk of the data?
- Which data are suspect, e.g. which data deviate from the appropriate standards provided that all data have been assigned to the correct trip level/depth which by itself is a crucial QC procedure of oceanographic data (Hood et al., 2010)
- Can the problem(s) be corrected? This recognizes that many problem data can be recovered either wholly or to a useful degree.

The first question involves checking the file trail for each parameter to see that the methodology was followed and in particular that standards were applied correctly. Assessment of the second and third issues is mostly a matter of identifying outliers and assessing their severity and cause. Especially in biological data it must be emphasized that suspicion of a data problem based on an outlying data value alone, without probable cause for an erroneous value, should normally not of itself be cause to ‘flag’ a value as questionable. Originally the term “outlier” denotes bad or incorrect data, but this need special treatment if applied to oceanographic data due to the real ocean variability. In biology what is a “spike” could well be a “signal”. Often, when one can see together all of the data, including calculated

parameters such as density, from a single station, the inter-relationships come clear between the various measured parameters of an unexpected or novel oceanographic feature. Some anomalies are real features of the ocean. Also, when stations are not spaced closely enough to resolve eddies, the eddies can be difficult to recognize. The “meeting” of water masses can result in considerable interleaving (though density compensated), seen not only on the CTD trace, but in chemical signatures also. Thus there is ample reason for exercising caution in rejecting ‘odd-looking’ data.

After examining many routine profiles, one tends to be lulled into thinking that it is ‘known’ what the ocean ‘should look like’, and then one too quickly flags unusual data points as questionable, or even omits them from the reported data. Outliers are cause for careful examination of data recording, standardizations, and computational correctness, but having done that, if no specific causes are found and the suspect value remains within the general realm of plausibility, it is usually best to let the value go as is. Any unusual features should be carefully checked in the original data and calculations. Anomalies in one property are likely to be reflected in other properties.

It is recommended to use the Flag scale developed under the Sea Data Net Project

<b>Flag Description</b>	<b>SEADATANET Flag</b>
no quality control	<b>0</b>
good value	<b>1</b>
probably good value	<b>2</b>
probably bad value	<b>3</b>
bad value	<b>4</b>
changed value	<b>5</b>
value below detection	<b>6 *</b>
value in excess	<b>7 *</b>
interpolated value	<b>8</b>
missing value	<b>9</b>
value phenomenon uncertain	<b>A</b>

\*these flags are not applicable to phytoplankton data

## **6. CONCLUSION**

As criteria for determining the acceptability of data from surveys of biological communities to meet specified information needs at international level are still under development the biological data QC will be based on:

- completeness of metadata will determine the quality/value of the data sets:
- professional judgement based on detailed template for the **metadata**

## 8. REFERENCES

Hood, E.M., C.L. Sabine, and B.M. Sloyan, eds. 2010. *The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines*. IOCCP Report Number 14, ICPO Publication Series Number 134. Available online at <http://www.go-ship.org/HydroMan.html>.

ICES Guidelines for Biological Plankton Data (Compiled August 2001 ; reviewed April 2006)

ICES. 2004. Biological monitoring: General guidelines for quality assurance. Ed. by H. Rees. ICES Techniques in Marine Environmental Sciences, No. 32. 44 pp.

ICES. 2005. Report of the ICES/HELCOM Steering Group on Quality Assurance of Biological Measurements in the Baltic Sea 148 pp.

ISO/IEC 17025. General requirements for the competence of testing and calibration laboratories and prEN 15204 (under Approval: 2005-11): Water quality -Guidance standard for the routine analysis of phytoplankton abundance and composition using inverted microscopy (Utermöhl technique).

Moncheva S., B. Parr, 2005 (updated 2010). Manual for Black sea phytoplankton sampling and analysis in the Black Sea.

REPORT OF THE ICES/HELCOM Steering Group on Quality Assurance of Biological Measurements in the Baltic Sea ICES Headquarters, Copenhagen, 25–28 February 2003, 69 pp

## LIST OF PARTICIPANTS

### Phytoplankton Workshop, Istanbul 21-23 June 2010

Country	Organisation	Expert Name	Contact details
Bulgaria	Institute of Oceanology-BAS, Varna	Assoc. Prof. Snejana Moncheva	Tel: +359 52 370485 Fax: : +359 52 370485 E-mail: <a href="mailto:snejanam@abv.bg">snejanam@abv.bg</a>
Romania	NIMRD – Constanta,	Laura Boichenko	Tel: Fax: : E-mail: <a href="mailto:laura_boicenco@cier.ro">laura_boicenco@cier.ro</a>
Russian Federation	IO P.P.Shirshov, RAS, Moscow	Dr. Alexander Mikaelyan	Tel: +7 499 124 59 74 Fax: +7 499 124 59 83 E-mail: <a href="mailto:mikaelyan@ocean.ru">mikaelyan@ocean.ru</a>
Turkey	Sinop University, Sinop	Mr. Fatih Sahin	Tel: Fax: : E-mail: <a href="mailto:fhshn@hotmail.com">fhshn@hotmail.com</a>
Ukraine	IBSS, Sebastopol	Mrs. Oleksandra Sergeeva	Tel: Fax: : E-mail: <a href="mailto:o.sergeyeva@ibss.org.ua">o.sergeyeva@ibss.org.ua</a>
Ukraine	IBSS, Sebastopol	Dr. Vladimir Vladimirov	Tel: Fax: : E-mail: <a href="mailto:v.vladymyrov@ibss.org.ua">v.vladymyrov@ibss.org.ua</a>
Ukraine	IBSS, Sebastopol	Dr. Yuliya Bryantseva	Tel: Fax: : E-mail: : <a href="mailto:brekall5@gmail.com">brekall5@gmail.com</a>
Ukraine	IBSS, Sebastopol	Mr. Denis Slipetsky	Tel: Fax: : E-mail: <a href="mailto:d.slipetsky@ibss.org.ua">d.slipetsky@ibss.org.ua</a>
Ukraine	Botanical Institute, Kiev	Dr. Alexander Krahmalnii	Tel: Fax: : E-mail: <a href="mailto:krakhmalnyy_a@mail.ru">krakhmalnyy_a@mail.ru</a>
Ukraine	Odessa University, Odessa	Mrs. Natalia Dereziuk	Tel: Fax: : E-mail: <a href="mailto:n.derezyuk@onu.edu.ua">n.derezyuk@onu.edu.ua</a>
	Black Sea Commission Permanent Secretariat	Prof. Ahmet Kideys	Tel: 0090 212 327 35 80 Fax: 0090 212 227 99 33 E-mail: <a href="mailto:ahmet.kideys@blacksea-commission.org">ahmet.kideys@blacksea-commission.org</a>

	Black Sea Commission Permanent Secretariat	Dr. Violeta Velikova	Tel: 0090 212 327 35 80 Fax: 0090 212 227 99 33 E-mail: <a href="mailto:violeta.velikova@blacksea-commission.org">violeta.velikova@blacksea-commission.org</a>
	Black Sea Commission Permanent Secretariat	Mr. Vladimir Myroshnychenko	Tel: 0090 212 327 35 80 Fax: 0090 212 227 99 33 E-mail: <a href="mailto:volodymyr.myroshnychenko@blacksea-commission.org">volodymyr.myroshnychenko@blacksea-commission.org</a>
	Black Sea Commission Permanent Secretariat	Ms. Nilufer Akpinar	Tel: 0090 212 327 35 80 Fax: 0090 212 227 99 33 E-mail: <a href="mailto:nilufer_akpinar@yahoo.com">nilufer_akpinar@yahoo.com</a>