

SOP 10 – Water-Quality Monitoring Data Reduction

Version 1.01 (June,2012)

The following table lists all changes that have been made to this Standard Operating Procedure since the original publication date. Any recommended or required changes added to the log must be complete and concise and promptly brought to the attention of the Project Leader. The Project Leader will review and incorporate all changes, officially complete the revision history log, and change the date and version number on the title page. For complete instructions, please refer to Instructions for Recording Revisions in SOP 13.

Revision History Log:

New Version #	Previous Version #	Revision Date	Author (full name, title, affiliation)	Location in Document and Description of Change	Reason for Change
1.01	1.0	6/1/2012	James M. Caldwell, Water Quality Specialist, USGS Maine Water Science Center	Updated all references to YSI dissolved oxygen probes with information specific to new optical sensor	Manufacturer's improvements

Spatial Water-Quality Surveys

Data collected by YSI sondes for the spatial component of water-quality monitoring is uploaded from the 650MDS data logger in PC6000 format, which results in a computer file with a .dat file extension. Upon installation, YSI EcoWatch assigns itself as the default program for PC6000 (.dat) files. Since EcoWatch does not allow editing of original data, this format is ideal for archiving purposes. In order to perform the data reduction steps in this SOP, however, it may be easier to create a second comma-delimited file (.cdf) that can be imported into a spreadsheet or database program. This is accomplished by using the <export> feature within EcoWatch. To assist in file tracking and management, change only the file extension when converting from one format to another, but do not alter the filename. Also create a text document (.txt) with this same filename for documenting all data reduction steps that are specific to that data file (date rejection). Data collected using LiCor quantum sensors for spatial monitoring of photosynthetically available radiation (PAR) are recorded on station data sheets (Appendix D). A back-up copy of the YSI data is also recorded on the field on the station data sheets, as is certain QA/QC data and information. For each station sampled during a spatial survey, perform

the data checks and data reduction steps described below. If any field or QC data are missing (not logged into memory), then data from the station data sheets should be manually substituted.

Dissolved Oxygen Calibration and Sulfide Interference Checks:

Check the calibration of the dissolved oxygen sensor that was logged immediately prior to each hydrocast. These in-air quality control data are logged to the data file and identified as site number 9. The in-air value for the % saturation of dissolved oxygen using the “local” pressure convention (DO Local %Sat) should read $\geq 98\%$ and $\leq 102\%$. If this is not the case, then the accuracy of the DO sensor may not be meeting data quality objectives and data from this hydrocast (all depths) should be omitted from subsequent data analyses. This occurrence should be very rare since the operator is instructed to perform a check in the field and take corrective action if necessary.

Sulfide Interference Check

The presence of hydrogen sulfide may interfere with the 6562 rapid pulse dissolved oxygen sensor used for dissolved oxygen and make the output “jumpy”. This effect is seen when the sensor encounters hydrogen sulfide in anoxic water or when it is inadvertently lowered into anoxic sediments. If a sonde touches the bottom during a deployment (recorded on the Station Data Sheet) or encounters anoxic water, subsequent DO data must be individually evaluated. Professional discretion may be used to reject data that may be influenced by sulfide interference.

Dissolved Oxygen Membrane Check

If a 6562 rapid pulse dissolved oxygen sensor was used check the “DO charge” for each reading (QC and all depths) to see that they fall within the range of 25-75. If any fall outside this range, discard the dissolved oxygen data (percent saturation and concentration) associated with this failed “DO charge” reading. The “DO charge” is not a feature on the 6150 ROX optical dissolved oxygen sensor.

Turbidity Interference Check

If a sonde touches the bottom during its deployment (reported on the Station Data Sheet), then bottom sediments were likely stirred up. These sediments will interfere with subsequent near-bottom turbidity readings, and may also interfere with chlorophyll *a* readings. Reject near-bottom turbidity and chlorophyll data from these casts.

Sensor Performance Specification Check

With the exception of some turbidity data (described below), all data that fall outside the design specification range for each sensor must be rejected. These specifications are currently as follows:

Temperature	-5 to 45 °C
Conductivity.....	0 to 100 mS/cm
Dissolved Oxygen, % saturation.....	0 to 500% air saturation
Dissolved Oxygen, mg/l	0 to 50 mg/l
Dissolved Oxygen charge	25-75 (6562 rapid pulse sensor only)
Salinity	0 to 70 ppt
Turbidity	0 to 1000 NTU
Chlorophyll	0 to 400 µg/l

Recent instrument performance evaluations conducted by the Alliance for Coastal Technologies (ACT 2006) showed abnormal negative turbidity readings using YSI 6-series sondes. Unlike some instruments, these sondes allow readings to go negative rather than restrain them at zero. Follow-up investigations by YSI, Inc. (ACT 2006b) have suggested that under typical conditions, zero turbidity standards may routinely become contaminated by the sonde, and that used turbidity probes will often generate some signal in zero turbidity solutions due to slight interferences in the optical patch from use. Because of this, small negative turbidity readings may be encountered that DO NOT need to be rejected from the record. When turbidity is found to go slightly negative, an offset may be applied to the entire set of turbidity readings gathered under that particular sonde calibration. The offset must be applied to all turbidity readings, and not just to the negative ones. Offsets up to two (2.0) NTU are permissible for this protocol.

Calculate the Average Station Location and Time

Average all of the latitude, longitude, and time data for each of the three depths at each station. Within the station data file, surface values are flagged with “site number” zero, mid-depth with 1, and near-bottom with 2. Be sure not to use quality control data collected during the in-air warm-up (station number 9) since the vessel may be underway during these periods.

Chlorophyll a

Chlorophyll *a* ($\mu\text{g l}^{-1}$) analysis is performed by extractive methods on near-surface water samples (see SOP 6 – Chlorophyll *a* Sampling and Analysis). Whether analyzed by the Atlantic Research Center at CACO or by an outside laboratory with intercalibration with the Cape Cod laboratory (see SOP 11 – Quality Assurance and Quality Control), these data should be added to the station data files – associated with the near-surface (site number zero) readings.

Processing of Light Data

Use the LiCor light data to calculate the attenuation coefficient of downwelling PAR (K_d).

- i. K_d should be first calculated as the slope of the least squares regression of $\ln(I_z/I_{air})$ against depth in meters, where I_z is the irradiance at depth z , and I_{air} is the irradiance in air. Units of K_d are m^{-1} .
- ii. I_z/I_{air} is recorded on the Station Data Sheet under the column labeled “Channel M2 - Underwater Irradiance- % of I_{Air} .” (Note: the data in this column are actually $I_z/I_{air} \times 100$, but this makes no difference.)
- iii. Calculate the coefficient of determination (r-squared) for the fit of this least square linear regression.
 - a. If the r-squared is 0.95 or greater, report the slope of this line as K_d for this site. This method accounts for variation in ambient irradiance caused by changes in cloud cover.
 - b. If the r-squared is less than 0.95, then calculate K_d as before, but substituting I_z for the ratio I_z/I_{air} . This data is labeled “underwater irradiance – PPFD” on the Station Data Sheet. If the r-squared for the new regression is 0.95 or greater, then use this new slope as K_d for the site. This method may yield better results if the readings in air (“deck” sensor) were faulty, but does not account for changes in cloud cover.
 - c. If neither approach yields a coefficient of determination ≥ 0.95 , then select the slope of the regression with the higher r-squared, provided it is ≥ 0.9 . If this condition cannot be met, then use best professional judgment to determine whether a representative K_d can be calculated by eliminating outliers in the data or including data that had originally been struck through by the field technician.

Continuous Water-Quality Monitoring

The purpose of this section of the SOP is to provide quality assurance guidelines for determining what portions of the continuous data should be included in the final record and used for analysis. Data rejection criteria are based upon recommendations and guidance from the sonde manufacturer, YSI, Inc. The following sections borrow liberally from data analysis guidance developed by YSI for the NOAA National Estuarine Research Reserve System for its System-Wide Monitoring Program (YSI 2000). By following closely with the NOAA NERR protocols, NPS will benefit from their years of experience in continuous estuarine monitoring as well as from the ability to interpret NPS data within a larger spatial context.

Data reduction for continuous monitoring is complicated by the fact that the sensors are deployed for an extended period between initial calibration and post-deployment calibration check. During this interval, sensor response may be corrupted by a number of factors including drift, biofouling, and catastrophic damage by organisms or extreme environmental conditions. The SOP for Continuous Water-Quality Monitoring is designed to maximize the collection of high-quality data that are accurate and representative of environmental conditions during the index

period. Drift and biofouling effects can usually be corrected using data from the post-deployment calibration check and the weekly quality assurance visits to the logging station (made during each of the spatial survey cruises at a minimum). Proper application of this SOP will assist in making data corrections where they are possible, and in culling out any remaining poor-quality data from the official record.

Data collected by YSI sondes for the continuous components of water-quality monitoring are uploaded from the sonde in PC6000 format with a .dat file extension. Upon installation, YSI EcoWatch software assigns itself as the default program for PC6000 (.dat) files. Since EcoWatch does not allow editing of original data, this format is ideal for archiving purposes. In order to perform the data reduction steps in this SOP, however, it is easier to create a second comma-delimited file (.cdf) that can be imported into a spreadsheet or database program. This can be accomplished either by uploading from the data logger a second time in the new format, or by using the <export> feature within EcoWatch. To assist in file tracking and management, change only the file extension when converting from one format to another, but do not alter the filename. Also create a text document (.txt) with this same filename for documenting all data reduction steps that are specific to that data file (data rejection).

Absolute Data Rejection

Absolute data rejection is a means of removing erroneous values from the data record because they violate basic principles of sensor behavior and/or proper sonde deployment. Because it is based upon basic principles, absolute data rejection can be accomplished using automated computer scripts with conditional arguments. There are at least two cases in which absolute data rejection must be invoked. The first of these is when the recorded value for any sensor falls outside its performance specifications. Environmental specifications for each sensor are listed in section Sensor Performance Specification Check above. Any data that fall outside the specification intervals should be removed from the official record.

The second case is when one or more of the sensors is believed to have come out of the water due to unexpected tidal fluctuations or improper deployment. In brackish and marine systems, this occurrence is indicated by anomalously low readings from the conductivity sensor. The conductivity sensor is the highest/shallowest of the sensors on the YSI 6600EDS and will be the first to be exposed to air. Since it is impossible to determine which other sensors were also out of water, all data from these out-of-water events should be removed from the record. Note that the upper PAR sensor from a YSI 6600EDS equipped with wiped PAR sensors will actually come out of the water before the conductivity probe, but this is difficult to discern using programming scripts and does not affect the other probes.

Discretionary Data Rejection, Drift Correction and Data Reduction.

This portion of the SOP is not easily automated using logic statements, and must be carried out manually under the supervision of a program manager with experience in continuous water-quality monitoring and knowledge of the YSI sonde and its application in this Protocol. Ideally, these steps should take place as soon as possible after sonde recovery so that the details of the

deployment are fresh in the minds of the field personnel, and so that corrective action can be attempted for any subsequent deployments. A rapid evaluation for any gross problems with the data record can be made using the graphing features of YSI's EcoWatch software. Examine the output from each sensor individually for any discontinuity in the data, which generally indicates catastrophic failure during the deployment. Check for battery voltage spikes in the data. Voltage spikes > 13.5 volts signal that one of the optical probes (turbidity, chlorophyll *a*, or ROX dissolved oxygen) may have malfunctioned. To check for sensor drift, examine the logged data in comparison with the discrete measurements that were made as part of the weekly spatial survey. Sensor-specific details for data rejection and drift correction are provided in the following sections.

Temperature:

The temperature sensor on the YSI 6560 temperature and conductivity probe is very robust and is unlikely to fail. Rare failure has been observed, however, and is most often associated with an irreversible malfunction due to leakage of the sensor housing. Incorrect temperature data are signaled by a clear point of temperature discontinuity followed by unreasonable and erratic values or unreasonable drift. If clearly incorrect temperature data are observed, all data from that point on should be removed from the official record. Since all the other sensors are temperature compensated using values from the 6560 thermistor, ALL data following failure of a temperature probe are suspect and should be removed from the official record.

Conductivity:

Like the thermistor on the YSI 6560 temperature and conductivity probe, the conductivity cell is very robust and rarely shows catastrophic failure. If an error occurs, it is usually a drift of the output due to biofouling within the water ports of the conductivity cell. This results in a change in the effective volume of the cell, which, in turn, results in drift of the output. The post-deployment calibration check will show whether such a drift has occurred. Cleaning of the sensor ports should resolve the issue. In these cases, at the discretion of the program manager, a linear compensation may be applied to the data using quality assurance data from the weekly surveys and from the post-deployment calibration check.

In the unlikely event of catastrophic sensor failure, a sharp discontinuity will be evident in the output, and all subsequent data should be removed from the official record.

Dissolved Oxygen:

The YSI 6562 dissolved oxygen probe is susceptible to both drift and catastrophic failure, although drift is less significant a problem when the sensor is mounted in the YSI 6600EDS. This is because drift usually results from biofouling over the probe membrane over the course of the deployment. The post-deployment calibration check will show whether such a drift has occurred. At the discretion of the program manager, a linear compensation may be applied to the data using quality assurance data from the weekly surveys and from the post-deployment calibration check.

Catastrophic failure of the dissolved oxygen probe is more common, and is usually the result of a puncture in the membrane, either by debris or organisms, or from an improperly installed membrane. Under these scenarios, output is generally characterized by a large discontinuity. In

most cases readings become unreasonably high very quickly and then either become very noisy or drift around. The likely cause of this behavior is “crosstalk” through the membrane hole caused by electrical continuity between the dissolved oxygen and conductivity sensors through the brackish estuarine water. All data after the initial discontinuity should be removed from the record, regardless of later probe behavior. After this type of catastrophic failure, the probe surface should be reconditioned before it is put back in service. Follow guidance in the YSI 6-Series Environmental Monitoring Systems Manual. Occasionally the DO probe will fail due to structural failure of the Clark-type electrodes. The symptoms are similar, and are accompanied by high DO-charge values. All data after the initial discontinuity should be removed from the official record.

The YSI 6150 ROX optical dissolved oxygen probe has excellent long term stability and should achieve $\pm 1\%$ accuracy when calibrated properly. However, physical obstructions from debris or a wiper failing to park properly may cause erroneous data. These types of interference often result in single large spikes in the data record. These values should be removed. Special attention should be paid to field notes for comments regarding wiper parking and sensor fouling.

Turbidity and Chlorophyll:

YSI’s optical turbidity and chlorophyll probes are very stable over time and not prone to drift. Consequently, there should be no need to manually adjust the data using linear compensation. When the turbidity and chlorophyll optical sensors produce erroneous data, it is usually due to physical obstructions in the measured volume of water. These obstructions could be from fouling organisms, debris, or wiper pads that fail to park in their proper locations. These types of interference are often characterized by very large negative values in the data record. These negative values should be removed from the record. Since they are outside the design specifications for the sensors, they will be removed automatically by the filters used for absolute data rejection (see section Absolute Data Rejection). Since it is possible for these physical obstructions to come and go over the course of a deployment, and since they generally cause no lasting effect on sensor response, it is NOT necessary to remove all subsequent data after one of these out-of-bounds data spikes has occurred. The program manager may use professional judgment in filtering out other optical data that are clearly inconsistent with the overall data record, such as occasional very-high positive peaks or high-frequency and very-high-amplitude noisy data (in which case the entire interval should be removed rather than just the peaks).

Photosynthetically Available Radiation:

The first step to assuring the quality of continuous PAR data is to compare pre and post deployment calibration data for each of the two LiCor 192SA sensors. These sensors are vulnerable to catastrophic failure from water vapor that slowly diffuses through the acrylic optical diffuser. If either sensor shows a dramatic drift ($>15\%$), then suspect catastrophic sensor failure. Through careful examination of the data record, it may be possible to identify a point when failure occurred. Remove from the final record all data after this point for the affected sensor only. Although PAR attenuation coefficients cannot be calculated during periods with only one functional sensor, it still provides a record of the underwater light field during the index period.

In some instances, catastrophic failure is not abrupt, making it more difficult to assign a cutoff point. The PAR data collected for quality assurance during the weekly spatial survey cruises are useful for identifying the timing of sensor failure, and also for applying a linear correction when PAR sensors have drifted less than 15%.

The principal metric of interest for this protocol is the attenuation coefficient of downwelling PAR, K_d . To calculate K_d , only a subset of the PAR data is used. This subset is all PAR data collected within 3 hours of local apparent noon. This is a slight extension to the interval recommended by Carruthers et al. (2001), but will still generate very good data while allowing comparison with data collected in the spatial survey. For all anticipated applications of the protocol, this time interval will run from 10:00 hrs to 16:00 hrs Eastern Daylight Savings Time. Include the data from the 10:00 AM logging event even though it technically represents the last minute of the preceding hour.

Create a new data column for K_d , and populate it (for the above subset) using:

$$K_d = \frac{\ln(I_z / I_0)}{-z}$$

where I_z is the reading from the lower PAR sensor, I_0 is the reading from the upper PAR sensor, and z is the separation between them in meters (0.5 m). Units for K_d are m^{-1} . Examine the logged K_d values against those collected for quality assurance purposes at the time of deployment, retrieval, and during each of the weekly spatial surveys. At the program manager's discretion, a correction factor and linear compensation for drift can be applied to the logged data using the quality-assurance data and the post-deployment calibration check.

Chlorophyll a Post-Correction

The chlorophyll sensor on the YSI 6600EDS is calibrated using a Rhodamine WT dye standard. Used alone, this calibration method will report consistently from deployment to deployment, but will not necessarily provide an accurate representation of actual pigment concentrations. More accurate reporting of chlorophyll *a* is achieved by calibrating these *in situ* fluorometric values against chlorophyll *a* data from extractive analyses on discrete grab samples. For each YSI sonde deployment, these calibration samples are collected at the beginning, end, and intermittently throughout each deployment (see SOP 6 – Chlorophyll *a* Sampling and Analysis).

Follow these steps to post-correct against extracted chlorophyll:

- i. Identify all the logged dye-calibrated chlorophyll samples that coincided with the collection of a discrete grab sample.
- ii. Perform a linear least squares regression of the dye-calibrated chlorophyll *a* data (dependent variable) against chlorophyll *a* from extractive analysis on discrete grab samples (independent variable).
- iii. Use the equation for this regression to correct chlorophyll *a* and create a new data field for corrected chlorophyll (retaining the uncorrected data as well).

Should the correlation coefficient for the regression be poor, or should the range of the correction data grossly under-represent field concentrations, then the program manager should use best professional judgment in deciding whether and how to apply the calibration. Data outside the range of the calibration samples may still be reported if professional experience suggests that the calibration would be reasonably accurate; however, these data should be flagged with uncertainty qualifiers in all databases and report products.

Data Management

A stipulation of the NPS Natural Resource Challenge (NRC) is that all physical, chemical, and biological water-quality data that are collected with NRC funds must be loaded into the NPS STORET database. Vital Signs Networks are encouraged to also archive any additional aquatic-related data collected with base Vital Signs funding in the NPS STORET. To facilitate this, the NPS Water Resources Division (WRD) has developed a series of MicroSoft Access®-based templates, called NPSTORET, that are patterned after the Natural Resource Data Templates. Each Vital Signs network must annually send its data to NPS WRD as an NPSTORET database or in its own database following a STORET Electronic Data Deliverable (NPSEDD) specification. NPS WRD will have responsibility for maintaining these data, and for transferring them into the U.S. Environmental Protection Agency's STORET National Warehouse. This is the ultimate archive of all NPS data related to physical, chemical, and biological water quality.

In concordance with the above requirements, data generated by this monitoring protocol will be managed by each network (NCBN and NETN) following guidance documented within their respective data and information management plans (Stevens and Entsminger 2004, Dieffenbach 2006).

References

- Alliance for Coastal Technology. 2006. Performance Verification Statement for the YSI 6600EDS Sonde and 6136 Turbidity Sensor. University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory. UMCES Technical Report Series: Ref. No. (UMCES) CBL 07-053
- Carruthers, T.J.B., B.J. Longstaff, W.C. Dennison, E.G. Abal and K. Aioi. 2001. Measurement of light penetration in relation to seagrass. In: F. Short and R. Coles (eds.) *Global Seagrass Research Methods*. Elsevier, Amsterdam.
- Dieffenbach, F. 2006. Data Management Plan. Northeast Temperate Network, National Park Service. Sept. 24, 2006. 50p.
- Stevens, S.M. and G. Entsminger. 2004. Information Management Plan. Northeast Coastal and Barrier Network Inventory and Monitoring Program. National Park Service. Draft: Dec 15, 2004. 99 p.
- YSI, 2000. Suggested Data Analysis Protocol for YSI 6000 Deployment Studies. Prepared for the National Oceanic and Atmospheric Association, National Estuarine Research Reserve System-Wide Monitoring Program.