AOP Workshop: Development of a Web-based Community AOP Processor

Vanessa M. Wright and Stanford B. Hooker (NASA CVO, Halethorpe, MD)

Hosted by University of California Santa Barbara (Santa Barbara, CA)
13-15 January 2009

The AOP Workshop was held at the University of California Santa Barbara from 13–15 January 2009, inclusive. The primary objective was to specify the requirements of a community maintained, open-source, Web-based interface for the Processing of Radiometric Observations of Seawater using Information Technologies, (PROSIT). Although it is envisioned that PROSIT will be capable of handling above- and in-water AOP measurements, this workshop was concerned with the latter. The structure of the workshop was as follows (the agenda appears at the end of this summary): oral presentations from each of the attendees on the first day of the meeting and then the remaining time was spent in working groups to define the required and desired specifications of a web-based processor, define how these specifications would impact the current ocean optics protocols, and to define the performance metrics needed to ensure data quality.

Prior NASA Data Analysis Round Robins (DARRs) have already established the variance imparted by processor-to-processor differences frequently equals or exceeds the total uncertainty budget permitted in calibration and validation field activities, which is no more than a few percent (currently about 3.5%). One way to remove this unwanted variability is to have the entire community use the same processor. To ensure easy access, a Web-based capability is desirable, and to ensure correctness and continuing improvement, an open-source architecture that the entire community can help maintain is preferred.

The working groups during the AOP workshop provided a framework for focusing on specific aspects of the processor and the required and desired components that fit into each category. There were six working groups with two further divided into subgroups as follows:

A Buoys
B1 Winch and Frame Profilers (including towed systems)
B2 Free-Fall Profilers
C1 Hyperspectral Sensors
C2 Fixed-wavelength Sensors
D Case-2 (Shallow) Waters
E Case-1 (Deep) Waters
F Performance Metrics

The principal objective of the working groups was to specify required and desired capabilities of the processor, and then to address how those capabilities will impact the current Ocean Optics Protocols, the performance metrics, and the terminology used to describe the quality of the data. A summary of a few of the key requirements of the processor associated with the different working groups is presented next.

The Buoy Working Group established the data processing steps that seemed to fit well when acquiring radiometric data on a buoy. This group supplied information in the form of a flow chart and demonstrated their need for a few more steps in the data processing scheme to allow for some scientific parameters including a spectral smoothing step and a time averaging step. Some requirements that were also suggested include quality checking and graphical display throughout parts of the processor and a manual section of data processing.

The Winch and Frame Working Group discussed some of the main issues associated with this type of deployment which included towed systems. The types of problems to be addressed included the incompleteness of legacy data sets (almost all of which were collected using winch and frame deployment systems), ship shadow, and self-shading. The requirements were an E_s sensor, description of the entire package, photo, location of deployment, ship heading, and sun data. The corrections that need to be applied include instrument self-shading and perturbation (requires IOP or chlorophyll data) and ship perturbation correction (includes shadow/reflection) and tilt sensors. The terminology for the quality categories was proposed to be as follows:

- Research (few or no corrections applied, tilt or solar reference data absent);
- Semi-quantitative (some corrections applied);
- Quantitative (suitable for calibration and validation activities, all standard corrections applied); and
- State-of-the-Art (suitable for vicarious calibration activities and next-generation problem sets).

The Free-Fall Working Group defined the required input levels to be instrument specific radiometric data, raw counts, and calibration data (including darks and immersion coefficients) and defined the desired inputs with capacity to apply multiple/time-averaged/interpolated calibrations, radiometric units and geophysical data. Potential inputs into the processor include but are not limited to: station data, instrument specification (including model, serial number, gain information, time, location, bottom depth information, etc.), CTD data, GPS data, and metadata (sky, sea, sun pictures, etc.). In the correction stage of the processor the required correction is depth data, pressure corrections, and sensor offsets and the desired corrections include: temperature effects, self-shading, E_s variation, wavelength normalization, and cosine correction. The quality/performance metrics were time from calibration, noise levels in E_d/L_u data as an indicator of bad K values, incorrect dark corrections, and sampling frequency.

The Case-1 Working Group was re-defined as clear water and vertically homogenous to avoid the case-1 and case-2 terminology. Raman scattering corrections and Fluorescence Line Height (FLH) are two important topics that came up in this working group. For now, the discussions were taken into consideration and they are listed as desired features for the processor. These topics are more for research questions and less for validation purposes. Along with this topic, FLH has not been researched well enough and more questions need to be answered and protocols need to be developed. Some questions include: what are the most effective wavelengths to determine FLH, and what bands should be put in the instruments to determine FLH? Additionally, at this time there has been no work done to validate the FLH protocols. However, it is a desired feature of the processor, especially the ability to process the FLH data to the research level.

The uncertainty budgets in the clear-water group are similar to uncertainties discussed above, but additional elements include: radiometric calibration uncertainty, instrument characterization, pressure calibration uncertainty, environmental conditions such as nonhomogenous conditions (e.g., wave focusing, fronts, sun angles, clouds) and biofouling. Some required uncertainties that were discussed in this group were the restart of SIRREX-like activity to assess calibration uncertainty and to establish the uncertainties on bio-optical models.

The Case-2 Working Group was re-defined as optically shallow with spatial and temporal variability, again to avoid using the case-1 and case-2 terminology. The required processor features include the bottom depth, type, and reflectivity, plus the ability to select multiple extrapolation intervals. The desired processor features include the bottom effect index and vertical resolution index. Performance metrics should be established for the bottom effect index and scale of spatial and temporal variation compared to scale of sampling (would need to be assessed by the PI).

There are four modules of the data processor: a) data ingestion, b) corrections, c) extrapolation intervals, and d) processing (which includes reprocessing). Four working groups were designated with the objective to determine the level of specificity of the data for each module.

The Data Ingestion Working Group concluded that the preferred format of the input files is ASCII. The required list of input data that denotes those necessary for the processor to function is as follows:

- Radiometric data (raw data so calibration and dark data is also needed)
- E_d, L_u, E_s, E_u
- Pressure (depth)
- Time and geolocation (longitude and latitude)

Below is the list of the inputs that are considered required for data to be included in the calibration and validation data set:

- Pitch/Roll
  - Profiling data and E_s
- Station Data
  - Cruise information (i.e., station, cast, and personnel)
  - Deployment conditions (i.e., freefall or winch/crane/wire)
    - Location (i.e., stern, rear quarter, etc.) and distance from ship
    - Boat/sun orientation with bottom information (i.e., depth, type)
- Instrument Data
  - Model, serial number, sensor dimensions, bandwidth and gain information
  - E_s sensor location and depth offsets
- Package description - include instrument layout and photos
- Field dark data

- Multiple cast information

Below is the list for the inputs that are the desired components of the processor:

- Temperature
- "Housekeeping" values (i.e., instrument temperature, voltages, etc.)
- Calibration Data
  - Lab dark scaling and offsets
  - Date, personnel, and lamp used during calibration and monitoring
- GPS Stream
- Comment field – transcription from logs
- Meteorological Data (including photos)
  - Wind, sea, and sky state; air temperature and ice conditions
  - Sun position
  - Aerosols
- CTD Data Stream*
- HPLC pigments and fluorometric chlorophylls*
- IOP Data*

* Note: for any of these ancillary data collected at a slightly different space or time, it is important to denote which or define how these profile or cast data are to be associated with the corresponding radiometric data.

The Corrections Working Group came up with a list of parameters that are required and desired for the processor. The first correction scheme is calibration. If the input data is raw instrument counts then required for calibration is the gain file, dark file and measurement equation to calibrate radiance/irradiance. Below is the list of the required and desired corrections to the calibrated data (designated as high, moderate, or low priority):

- Time Synching - High
- Pressure tare - High
  - Surface file
- Depth correction for radiance / irradiance to pressure sensor – High
  - Position information
- Self-shading - High
  - Need algorithm for each instrument
  - Size / shape of package
  - Diffuse to direct sky / cloud (relationship to solar geometry)
  - IOPs or path to get them from AOPs
- Wavelength co-registration – High
  - Choice of algorithm for interpolation
  - Bandpass differences between instruments
- E_s variation – High
  - Option of normalizing profile E_d/L_u with E_s observation
- Cosine collector correction - Moderate
  - Need cosine response curve and radiance distribution and tilt
- Immersion coefficients* - High
  - Radiance - changes of field of view

        Irradiance - less importance for cal/val because E_s is used in nL_w
    Platform perturbations (ship, tower, bridle, etc.)* - Low
* Need uncertainty bounds and requires research

Below is the list of the required filtering operations associated with the correction module of the processor:

- Tilt – High
    - Mask / flag data for set range values
- De-spike – High
- Clouds – High
    - Mask / flag for highly variable clouds
- Low signal levels – High
    - Set noise equivalent radiance/irradiance levels

Below is the list of the desired corrections:

- Raman
- Polarization
- Exact "L_w" and true BRDF correction
- Mismatch of time constants among measurement suite on same package
- Hysteresis
- Biofouling
- Bubbles
- Bioluminescence
- Uncertainty of all of these corrections

The Extrapolation Intervals Working Group came up with a set of criteria for the processor to use in determining the data values used for near-surface AOP extrapolation. The first criteria is the tilt criterion of $< 5°$ for non-overcast skies (for overcast skies it can be relaxed to 8°) for satellite cal/val suitable measurements. Greater tilt values would be put into the research-suitable quality category. The next criterion for selecting the extrapolation interval is the homogenous layer criteria. The temperature profile data will be used to establish the depth of the homogenous layer. The extrapolation interval should be established in the red wavelengths first, then may extrapolate from a deeper limit for the blue wavelengths as long as the interval remains in the homogenous layer.

The same extrapolation interval is to be used for all radiometric quantities (i.e., L_u, E_d, E_u). The goal is to strive for an automated choice of the interval, but an option should still exist for manual override, with mandatory explanation. Some metrics when selecting the extrapolation interval include the variance of E_s, tilt, and temperature within the layer, number of sampling points used for the extrapolation, and to use sky conditions to notify user that wave focusing may be a problem and the shallowest depth for the extrapolation may not be from the shallowest data (alert for manual user examination).

The Processing Working Group came up with a list of steps to be put into the processor:

1) Separate processing for fixed depth and profiling instruments
2) Base calculations (profile, time flag) - error calculations and data flags

3) Optional binning - select bin widths (time and depth)
4) Surface products calculations - error calculations (interactive aggregation)
5) Multi-cast aggregation and selection
6) Calculate statistics and uncertainties
7) Output / Archive - selected outputs, three files (profile, binned, bulk data)
8) Format and prepare for submission to SeaBASS
9) Processing lineage / database for reprocessing facilitation

Below is the list of the surface products that were noted in the Processing group:

L_u, E_d, E_u $(\lambda, 0^{-,+})$
K_d and K_l $(\lambda, \Delta z)$ - extrapolated, use to get L_u and E_s
$nL_w(\lambda)$, $L_w(\lambda)$, and “exact” $nL_w(\lambda)$
$R(\lambda)$ (irradiance) and $R_{rs}(\lambda)$ (radiance)
$E_s(\lambda)$ – average, $E_{sky}(\lambda)$, $E_{sky}/E_s(\lambda)$
$Q(\lambda)$
OC algorithm ratios

Below is the list of the bulk products that were noted in the Processing group:

K_d $(\lambda, 0^-$ to $1/e^*E_s)$ or for 1% light level or Muellers 37% light level
$K_{par}(\lambda)$
Z_{par} (1/e), Z_{par} 10%, and Z_{par} 1%
FLH
Uncertainties

Future Tasks that were discussed included establishing uncertainty budgets tests will be done on intercomparing the community Web-based processor, environmental variability, repetitive casts, replicates, independently check with another group, and self-shading issues (get information from a few different water types).

Important changes to the protocols were discussed and some of these changes include: extrapolation interval selection, emphasis to be placed on more than one cast (collect redundant data to assess variability), deep casts first and then many short casts, ensure proper documentation of where the solar reference sensor is located, and distance from the ship (20–30 m is recommended, however the literature says at least 9 m but this depends on the size of the ship and a number of other variables).

The next step will be to begin implementing the required features of the processor into the existing CVO processor to develop the community Web-based AOP processor (many of the required capabilities are already incorporated). The CVO processor has already been ported to a Web-based application and can be used over the internet or on a single computer (a desired capability, so the processor can be taken into the field where internet access is usually problematic). Also, scientists will be visiting the CVO to discuss the processor and view a demonstration. Overall, the scientific community was responsive to the AOP processor and agrees that there is a strong requirement for a community processor with established performance metrics and uncertainties in the data products.

Workshop Attendees

AOP Workshop, Hosted by University of California Santa Barbara (Santa Barbara, CA) January 2009		
Help and Support:	Location:	E-mail:
Kris Duckett	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	kris@icess.ucsb.edu
Kathy Scheidermen	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	kathys@icess.ucsb.edu
Attendees:		
Richard Zimmerman	Old Dominion University, Norfolk, Virgina, USA	rzimmerm@odu.edu
Dave Siegel	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	davey@icess.ucsd.edu
David English	University of South Florida, St. Petersburg, Florida, USA	denglish@marine.usf.edu
David Antoine	Laboratoire d'Oceanographie de Villefranche, Villefranch-sur-Mer, France	antoine@obs-vlfr.fr
Jeremy Werdell	NASA Goddard Space Flight Center, Maryland, USA	jeremy.werdell@nasa.gov
Sean Bailey	NASA Goddard Space Flight Center, Maryland, USA	sean.w.bailey@nasa.gov
Mati Kahru	Scripps Institution of Oceanography, UCSD, San Diego, USA	mkahru@ucsd.edu
Dariusz Stramski	Scripps Institution of Oceanography, UCSD, San Diego, USA	dstramski@ucsd.edu
John Morrow	Biospherical Instruments Inc., San Diego, California, USA	morrow@biospherical.com
Germar Bernhard	Biospherical Instruments Inc., San Diego, California, USA	bernhard@biospherical.com
Heidi Sosik	Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA	hsosik@whoi.edu
David Dana	HOBI Labs Inc. Bellevue, Washington, USA	dana@hobilabs.com
David Court	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	dcourt@icess.usbc.edu
Stephane Maritorena	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	stephane@icess.ucsb.edu
Norm Nelson	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	norm@icess.ucsd.edu
David Menzies	Institute for Computation Earth System Science, UCSB, Santa Barbara, USA	davem@icess.ucsb.edu
Carlos Alberto Garcia	Federal University of Rio Grande, Brazil	dfsgar@furg.br
Cara Wilson	Southwest Fisheries Science Center, NOAA, Pacific Grove, California, USA	cara.wilson@noaa.gov
Stephanie Flora	Moss Landing Marine Laboratories, Moss Landing, California, USA	flora@mlml.calstate.edu
Wendy Kozlowski	Scripps Institution of Oceanography, UCSD, San Diego, USA	wkoz@ucsd.edu
Carol Johnson	NIST, Gaithersburg, Maryland, USA	cjohnson@nist.gov
Vanessa Wright	NASA/CVO Halethorpe, Maryland, USA	vanessa.m.wright@nasa.gov
Stan Hooker	NASA/CVO Halethorpe, Maryland, USA	stanford.b.hooker@nasa.gov

Workshop Agenda

Time	13 January (Tue)	14 January (Wed)	15 January (Thu)
0830	*Welcome*	*Welcome*	*Welcome*
0845	Workshop Introduction (Hooker)	Legacy Processors (Siegel)	Working Group Report A and Discussion
0900		SeaBASS Lessons (Werdell)	
0915	Survey Summary (Hooker)		Working Group Report B and Discussion
0930	Web-Based Processor (Hooker)	Practical Aspects of Calibration and Validation Quality (Bailey)	
0945			Working Group Report C and Discussion
1000	GSFC Processor (Hooker)	Hyperspectral Processing and Case-2 Considerations (Dana)	
1015	ODU Processor (Zimmerman)		
1030	*Break*	*Break*	*Break*
1100	UCSB Processor (Siegel)	Working Groups (A and B)	Working Group Report D and Discussion
1115	USF Processor (English)		
1130	LOV Processor (Antoine)		Working Group Report E and Discussion
1145			
1200	SeaBASS Processor (Werdell and Bailey)		Working Group Report F and Discussion
1215			
1230	*Lunch*	*Lunch*	*Lunch*
1300			
1330			
1400	Scripps Processor (Kahru)	Working Groups (C and D)	Plenary Discussions: Required and Desired Features of a Web-based Community Processor, Changes to the Ocean Optics Protocols, and Performance Metrics
1415	Biospherical Processor (Morrow)		
1430			
1445	Scripps Processor (Stramski)		
1500	HOBI Labs Processor (Dana)		
1515			
1530	*Break*	*Break*	*Break*
1600	WHOI Processor (Sosik)	Working Groups (E and F)	Plenary Discussions: Workshop Report and Writing Assignments
1615	FURG Processor (Garcia)		
1630	NOAA Processor (Wilson)		
1645	Scripps Processor (Kozlowski)		
1700	MLML Processor (Flora and Johnson)		
1715			
1730	*Adjourn*	*Adjourn*	*Adjourn*