Uncertainties of in situ Ocean Color Radiometric Measurements

Carol Johnson
NIST

Outline

- Why bother?
- Some technical background
- Approach to follow
- In situ radiometric ocean color data

Why Bother?

Traceability

"Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons, all having stated uncertainties." (VIM)

Completeness

The results of measurements are estimates or approximations to the value of the measurand and thus complete only when accompanied by a statement of uncertainty.

Validation by Comparisons

Absolute differences in results, e.g. from independent methods, instruments, or laboratories, *in spite of what some publications may infer are not related in any fashion* to the uncertainty in each participant's result. Scientific evaluation of independent results relies on comparisons of results in the context of their respective uncertainties.

We Learn Something

Examples: How to allocate resources, where to spend time characterizing the instruments, how to improve the experiment design, and hopefully, something about the "truth".

Guidelines Exist

- http://www.bipm.org/en/publications/guides/
  - GUM: Guide to the Expression of Uncertainty in Measurement
  - VIM: International Vocabulary of Metrology

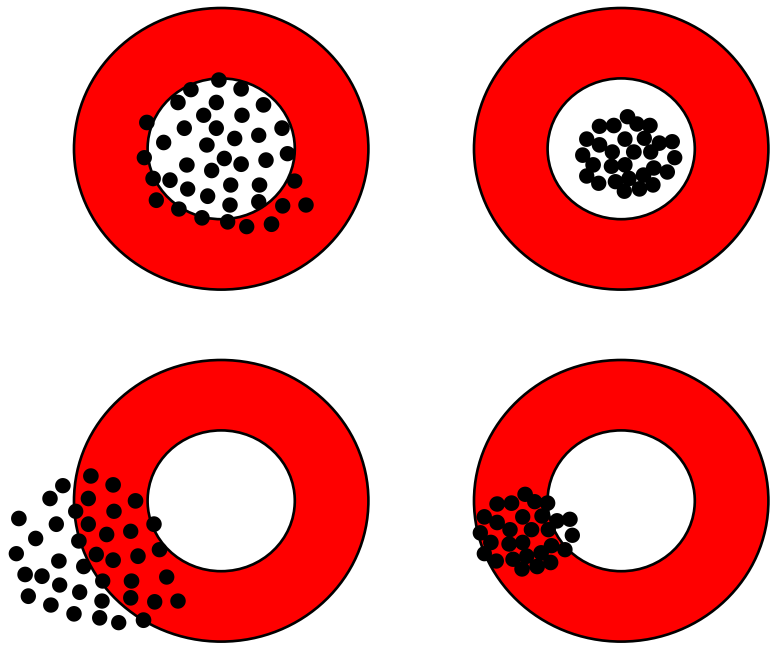
Uncertainty of measurement is a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

- Uncertainties are data products; they are assessed quantitatively; they should be part of the data management structure.

Accuracy and Precision

Accuracy is a qualitative concept regarding the closeness of the agreement between the result of a measurement and the true value of the measurand

Precision is a qualitative concept regarding self consistency of the results of a measurement



Imprecise & Accurate

Precise & Accurate

Imprecise & Inaccurate

Precise & Inaccurate

Accuracy & Precision as Requirements

The required accuracy of the measurement is necessary to determine the specification and definition for the measurand; it should be defined with sufficient completeness so that for all practical purposes its value is unique (GUM, 3.1.3).

Example

> “The requirement for ocean color vic/cal is 5% in situ L_w.”
>
> “The requirement for ocean color vic/cal is 5% relative standard uncertainty at 490 nm in L_w for observations of clear, open oceanic waters with stable marine atmospheres.”

The required precision of the measurement is determined in context of the magnitude of uncertainty components arising from random variability compared to those arising from systematic effects.

Example

> An integrating sphere source with short term (1 min) stability of 1 part in 50,000 and drift of < 0.1%/hour is compatible with an 1% accuracy goal.

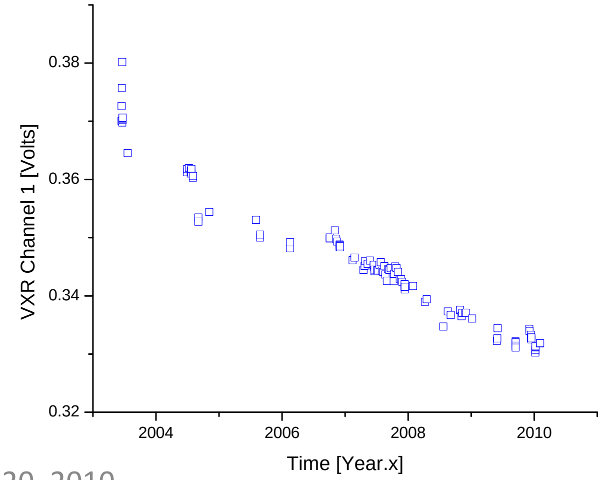
Error vs. Uncertainty

Error of the result of a measurement and the true value of the measurand are both unknowable; uncertainties are quantities we estimate using data or other information.

Random error arises from unpredictable variations (random effects), given rise to variability in the results of repeat measurements. “Result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions.”

Systematic error cannot be eliminated but it often can be reduced by eliminating systematic effects. “Mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus the true value of the measurand.”

Example: Filter radiometer measurements of an integrating sphere source: Random error is contributing to the scatter and systematic error, here a time dependent degradation of the filter radiometer, is causing the trend.



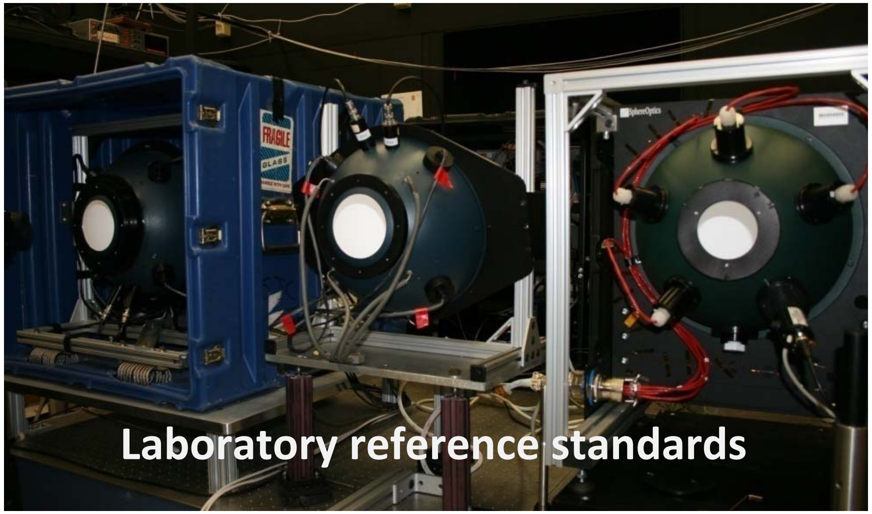
Repeatability vs. Reproducibility

Random error -- generally speaking, to reduce, make additional measurements under conditions of Repeatability

Systematic error -- generally speaking, to reduce, characterize the instrumentation and assess/validate under conditions of Reproducibility

Repeatability (Reproducibility) – same (different) conditions of measurements for the same quantity: e.g., principle, method, observer, location, instrument, time.

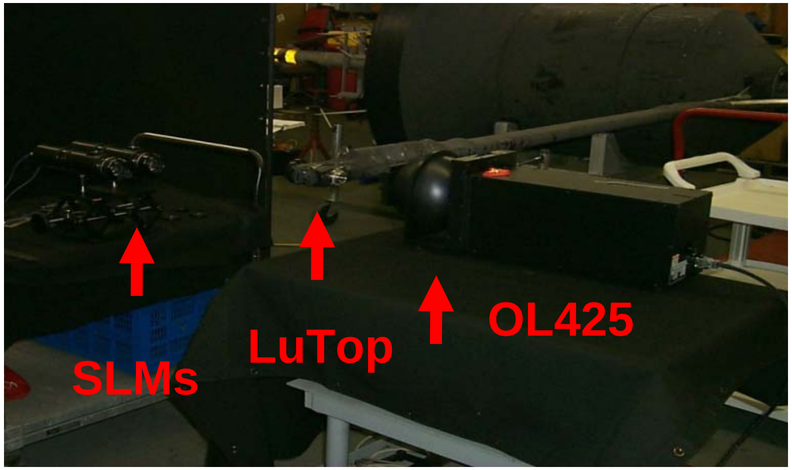
Our problem is we have no control of the conditions of measurement for most natural environments, requiring robust and redundant experimental design.



Example – Spectral Radiance

<u>Repeatability</u>

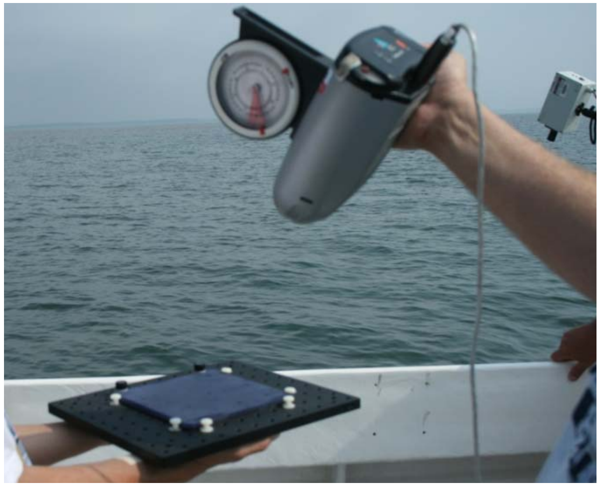
Measurand – LuTop responsivity [Counts/L(λ)]
Principle – Measure radiometric reference (OL425)
Location – MOBY facility, Snug Harbor Honolulu
Observer – Michael Feinholz
Instrument – B2xx, LuTop
Time – ~ 20 min, darks, lights, darks
Ancillary data – Ambient temperature for responsivity(temperature correction), Independent radiometers (SLMs) for OL425 drift correction, stray light correction algorithm



Example, Spectral Reflectance

<u>Reproducibility</u>

Measurand – Hemispherical/directional spectral reflectance [Unitless]
Principle – Incorporate stable unknown reference target into In Air Workshop
Location – Long Island Sound (Asharoken, New York)
Observers – CUNY, NRL/SSC, MSU/SSC, NOS/NOAA, NESDIS/NOAA
Instruments– NIST ground glass color tile standard; Two ASDs, Two Spectrix', HyperPRO, and GER spectroradiometers
Time – Sequential tests during the day's activities (August 2010)
Ancillary data – Varied with observer (time, lat/long, sky conditions, etc.)



More Terminology

Standard uncertainty is the uncertainty of the result of a measurement expressed as a standard deviation. Combined standard uncertainty – corresponds to the overall, complete estimate of standard uncertainty from multiple sources.

Coverage factor is a multiplier for the standard uncertainty; the expanded uncertainty defines an interval that encompasses a large fraction of expected results for the determination of the measurand.

Type A method of evaluation of uncertainty is by statistical analysis of observations; Type B is by means other than statistical.

Corrections and Correction Factors remove bias from systematic effects and they have associated uncertainties:

A correction is a value added algebraically to the uncorrected result of a measurement to compensate for systematic error.

A correction factor is a numerical factor by which the uncorrected result of a measurement is multiplied to compensate for systematic error.

Evaluating Standard Uncertainty

$$S = G\int L(\lambda)\,R(\lambda)\,d\lambda$$

Filter radiometer for radiance
G=preamp gain, *L*=spectral radiance of source, *R*=absolute system level radiance responsivty, λ=wavelength, *S*=net signal

Model the measurement

Evaluate the input quantities (*G*, *L*, *R*)

Determine their uncertainties

Calculate the measurand (*S*)

Determine the combined standard uncertainty (requires information on any correlation among the input quantities as well as a complete list of contributors)

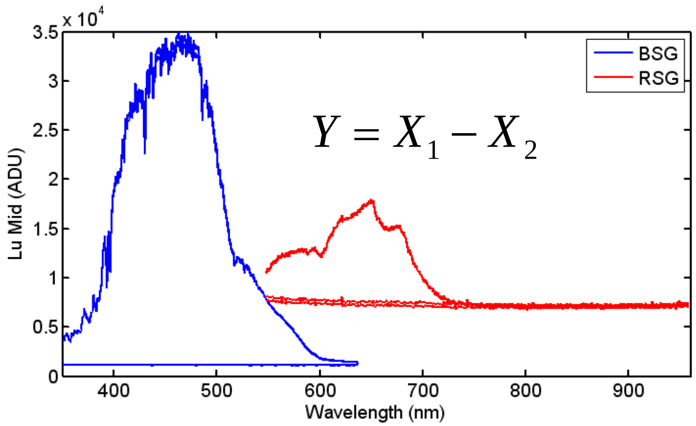
Calculate the expanded uncertainty

Report the value and the standard or expanded uncertainty

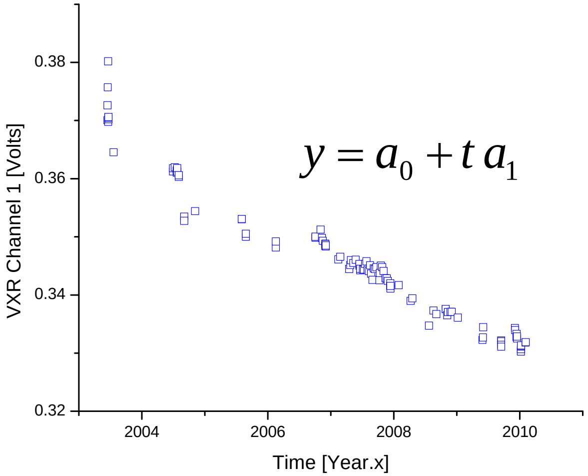
Uncertainty Tables should include designation on each component as Type A (with number of degrees of freedom) or Type B

Model the Measurement

Using analytical functions = measurement equation(s): $y = f(x_1, x_2, \dots x_M)$



Net signal: subtract shuttered (dark) from open (light)



Interpolation: System response at time *t*

Determine sensitivity coefficients for *f* experimentally if the model is not analytical

Example: What is the change in $L(\lambda)$ with lamp operating current for an integrating sphere source?
Answer – characterize the source using a stable spectroradiometer.

Determine sensitivity coefficients for *f* numerically using an algorithm

Example: What is uncertainty component in the immersion coefficient for spectral radiance due to the natural variability in oceanic salinity and temperature?
Answer – evaluate the impact of changes in the index of refraction for the range of input values.

Determining Uncertainty Values - A

Type A values are based on a frequency distribution. Type A does not imply the contributions are from random effects, it designates the method of evaluation ONLY.

I. Find the experimental standard deviation of the mean from *n* independent, multiple, random measurements q_{kC}

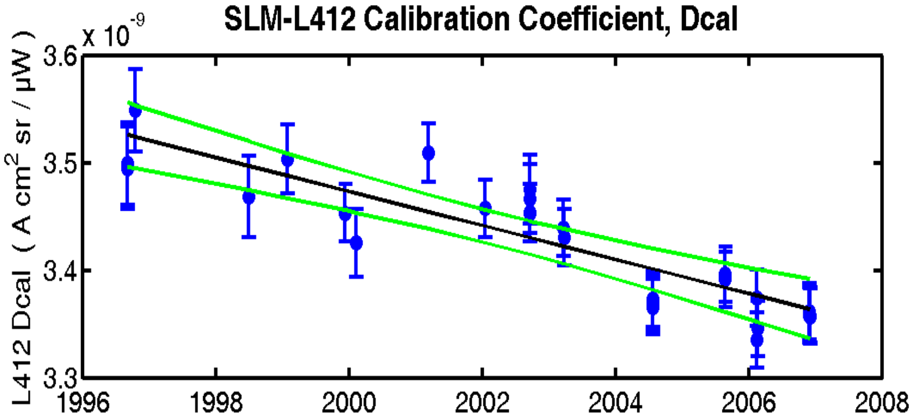
$$u^2(\bar{q}) = \frac{1}{n}\frac{1}{n-1}\sum_{k=1}^{n}\left(q_k - \bar{q}\right)^2 \qquad \bar{q} = \frac{1}{n}\sum_{k=1}^{n} q_k$$

II. When interpolating in fitted data, use established statistical procedures to get variances and covariances of the fitted parameters, then apply using same principles as for determining combined standard uncertainty.

If the model is linear $y = a_0 + t\,a_1$

$$u^2(y) = u^2(a_0) + t^2 u^2(a_1) + 2t\,u(a_0)\,u(a_1)\,r(a_0, a_1)$$

$r\,(a_0, a_1)$ = correlation coefficient



Determining Uncertainty Values - B

Type B values are based on a priori distributions. Type B does not imply the contributions are from systematic effects, it designates the method of evaluation ONLY.

I. Somebody tells you (e.g., a calibration report from a supplier)

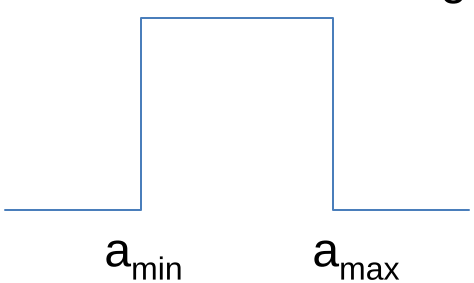
"Land-level" integrating sphere source, expanded uncertainties in spectral radiance

Source of Uncertainty	Relative Expanded Uncertainties [%] (*k*=2)							
	300	**400**	**500**	**600**	**700**	**800**	**900**	**1000**
Blackbody quality	0.12	0.07	0.03	0.01	0.00	0.01	0.03	0.04
Calibration of pyrometer lamp	0.33	0.27	0.22	0.18	0.15	0.12	0.11	0.10
BB temperature determination and transfer to lamp	1.02	0.21	0.36	0.16	0.37	0.31	0.31	0.55
Wavelength accuracy	0.12	0.10	0.07	0.06	0.05	0.04	0.04	0.03
Temperature scale (thermodynamic vs ITS-90)	0.58	0.46	0.37	0.30	0.27	0.24	0.20	0.19
Unc. for NPR (4 lamps) (*k*=2)	**1.23**	**0.59**	**0.57**	**0.39**	**0.48**	**0.41**	**0.39**	**0.59**

II. Scientific judgment: "There is a 50-50 chance the value lies between a_{min} and a_{max}."

$$u(x) = 1.48\,a, \quad \text{where } a = \left(a_{\max} - a_{\min}\right)/2$$

III. Reasonable knowledge of the minimum and maximum values for the measurand:



$$u^2(x_i) = \frac{\left(a_{\max} - a_{\min}\right)^2}{12}$$

Uniform probability distribution

Combined Standard Uncertainty

Measurement model $y = f(x_1, x_2, \ldots x_M); \partial f / \partial x_i =$ sensitivity coefficients

Independent quantities x_i, no correlation, u_c is combined standard uncertainty in *y*

$$u_c^2(y) = \sum_{i=1}^{M} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i)$$

Correlation among the quantities; $r(x_i, x_j)$ = correlation coefficients

$$u_c^2(y) = \sum_{i=1}^{M} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i) u(x_j) r(x_i, x_j)$$

Quantities are 100% correlated

$$u_c^2(y) = \left[\sum_{i=1}^{M} \frac{\partial f}{\partial x_i} u(x_i) \right]^2$$

Correlation

Uncorrelated – input quantities are independent; off-diagonal elements in the covariance (or correlation coefficient matrix) are zero.

Correlated – input quantities are interdependent, a change in one implies a change in the other; off-diagonal elements in the covariance matrix are not zero.

We make *n* independent pairs of simultaneous observations of two random variables in order to assess their covariance (C.3.4 in GUM).

Example: Calibration of several integrating sphere sources of spectral radiance using a common reference standard. (We assume close in time so as to be effectively simultaneous.) The estimated values for $L(\lambda)$ are correlated to those of the reference standard. The degree of correlation depends on the ratio of the uncertainty of the comparison (the transfer uncertainty) to the uncertainty of the standard. If the uncertainty of the comparison is negligible, the uncertainty in $L(\lambda)$ is equal to $u(L(\lambda)_{ref})$ – there is 100% correlation.

Ocean Color in situ & Simultaneity

Typically, measurements for different input observations for the same instrument/method are performed sequential in time – not simultaneous, e.g.,

In-water radiometry – $L_u(z,\lambda,t)$, $E_s(\lambda,t)$ to derive $K_L(z,\lambda)$ and $L_w(\lambda)$

In-air radiometry – $L_{surf}(\lambda,t)$, $L_{sky}(\lambda,t)$, to derive $L_w(\lambda)$ with calibrated radiometers

In-air radiometry – $L_{surf}(\lambda,t)$, $L_{sky}(\lambda,t)$, $L_{ref\ target}(\lambda,t)$, to derive remote sensing reflectance with uncalibrated radiometers

What might cause correlation among these different random variables?
Sun's position; Clouds
Other sky conditions (e.g, thin cirrus or variable atmosphere)
Adjacency effects – moving structures in the near field
Natural (rapid) variability
Sea state (bubbles, effects from wind, waves)
Reference standards (e.g., diffuse reflectance target)
Wave focusing
Depth, including effect of tilt
Tilt (view angle) through BRDF of radiance distributions

Thoughts to Keep in Mind, 1

The GUM assumes the measurand had a unique value. The GUM takes no position on how to draw conclusions from results of comparisons. (But see information on Key Comparison Reference Values and Degrees of Equivalence.)

http://kcdb.bipm.org/default.asp

What uncertainty components are really associated with random effects? That is, do I expect to get a better estimate if I perform additional measurements? Carefully identify which components contribute to variability and which do not.

G. Fraser, C. Gibson, H. Yoon, and A. Parr, "'Once is Enough' in Radiometric Calibrations," *J. Res.NIST*, **117**, 41-51 (2007).

Pay attention. Under conditions of repeatability, is the experimental standard deviation consistent with what you'd expect from the uncertainty budget?

Redundancy (and thorough characterization) are good approaches to sanity – internal monitor detectors on all sources, internal sources on all radiometers….

Thoughts to Keep in Mind, 2

Establish, as much as possible, control charts, check standards, and data quality indicators as part of the routine protocols. Rigorously adhere to defined practices.

Do retrieved $E_s(\lambda)$ values agree with clear-sky irradiance models ?

Are retrieved $K_L(\lambda,z)$ > clear water values and self-consistent ?

Did the dominant CIE wavelength change?

Unrecognized systematic effects cannot be accounted for in the uncertainty budget but they do contribute to the error of the measurement.

Perform blind comparisons

Don’t Double-Count. Effects that are considered in Type B evaluations may be contributing to variability that is already accounted for in the Type A evaluation of another component, requiring reconsideration of the Type B value.

Example: sensitivity of the alignment of radiometer to the reference source is evaluated in a dedicated characterization experiment and a Type B with uniform distribution evaluated. If, however, realignment is a normal part of the process, its effect will be captured in a Type A evaluation (e.g., repeats of radiance responsivity).

Vicarious Calibration for Ocean Color Satellite Sensors

Water-leaving spectral radiance is a small fraction of the top-of-the atmosphere spectral radiance.

$$L_{t}(\lambda) = \left[L_{r}(\lambda) + L_{a}(\lambda) + t_{d_{v}}(\lambda)\left(L_{f}(\lambda) + L_{w}(\lambda)\right)\right] t_{g_{v}}(\lambda)\, t_{g_{s}}(\lambda)\; f_{p}(\lambda)$$

The relative standard uncertainty in $L_t(\lambda)$ is stringent, ~ 0.2%

Global Observations: $L_{t}(\lambda) \rightarrow L_{a}(\lambda) \rightarrow L_{w}(\lambda) \rightarrow L_{wn}(\lambda) \rightarrow$ Bio-optical products

Vicarious Calibration (applied to satellite beginning of life gains):

$$L_{w}^{r}(\lambda) \rightarrow L_{wn}^{r}(\lambda) \rightarrow L_{t}^{r}(\lambda) \rightarrow \frac{L_{t}^{r}(\lambda)}{L_{t}(\lambda)} =$$ Gain correction factors

MOBY Uncertainty Budget

Uncertainty Component	Method of Evaluation Uncertainty	Type	Causes Variability?
Radiance Reference (ISS)	NIST Cal. Rpt.	B	No
Drift over lamp life (50h)	beg/end of life NIST calibrations	B	No
Interpolation in wavelength	Compared techniques	B	No
Reproducibility	Different cal. source same buoy	B	No
Wavelength accuracy	Fitted uncertainty and dL/dl	B	No
Stray light (w correction)	Monte Carlo studies	B	No (1)
Temperature (w correction)	Uniform probability	B	No
Immersion coefficient	Uniform probability	B	No (2)
Stability during deployment	Pre/post cal. Differences + ESD	B	No (2)
Internal cal. sources	ESD	A	No
Wavelength stability	Fraunhoffer lines in Es	B	No
Exp. std. dev. (ESD)	GUM Eq (5) in 4.2.3	A	Yes
Temporal overlap	Uniform probability	B	Yes
Self shading	10% of modeled effect	B	No
Biofouling	Diver cal. lamps	B	No
Arm separation	Uniform probability	B	No
Depth, top arm (w tilt)	Uniform probability	B	No (3)
Cosine response Es	Apply fitted results w tilts	A	No (3)
Lu distribution	BRDF data, tilts, modeling	B	No (3)
Extrapolation to/thru surface	High depth resolution data sets & modeling	B	No

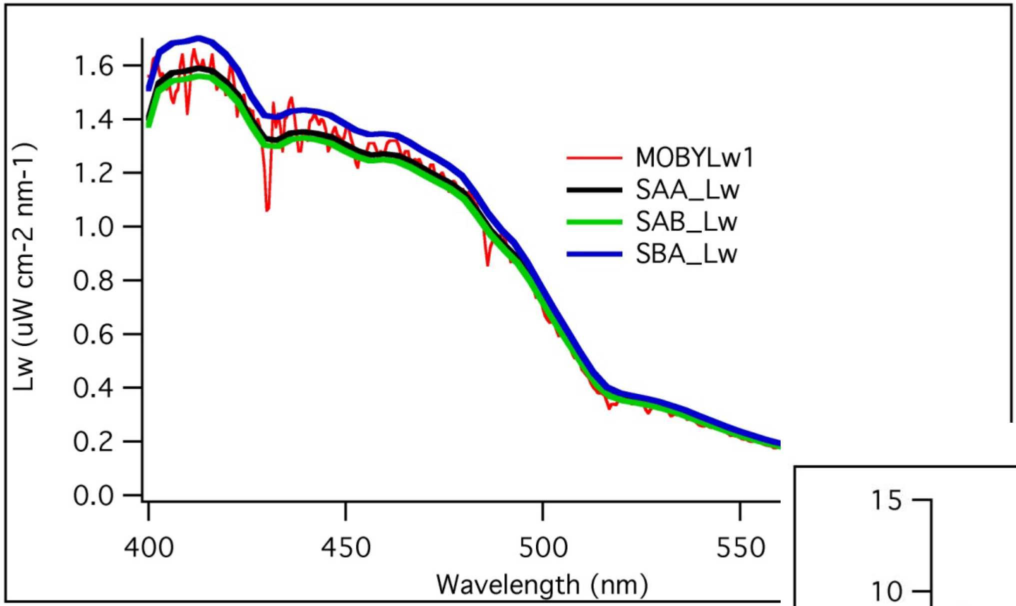
(1) There is long term variability – epochs in stray light performance; (2) Not likely to be variable on a daily basis; (3) We assume here the tilt is ~constant over one hour file collect. This is a work in progress; not every task (component) is complete.

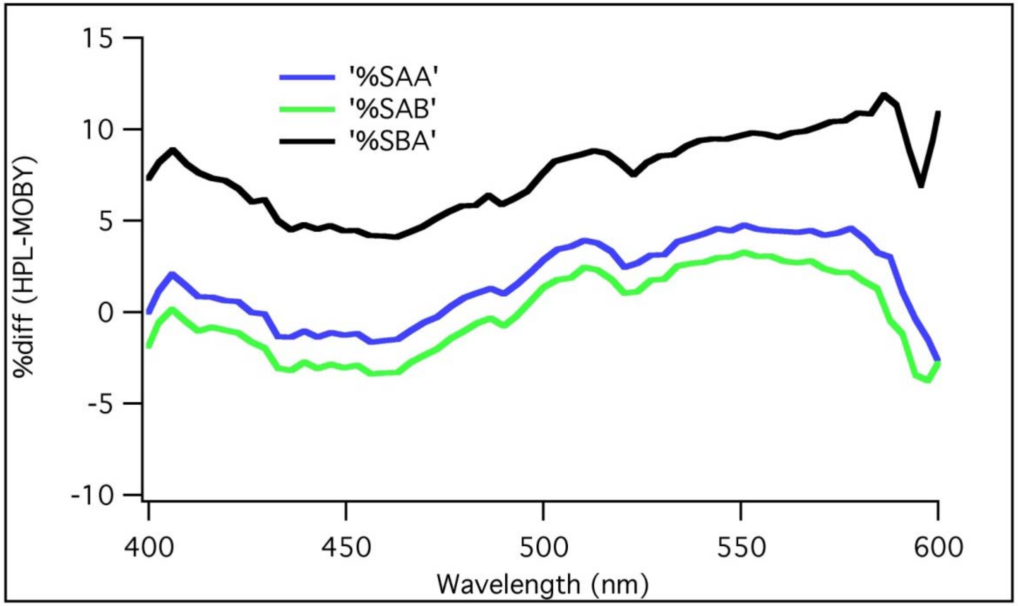
Uncertainty Estimates– $L_u$top B231

Uncertainty Component [%]	8 411.8 nm	9 442.1 nm	10 486.9 nm	11 529.7 nm	12 546.8 nm	13 665.6 nm
Radiometric Calibration Source						
Spectral radiance	0.65	0.60	0.53	0.47	0.45	0.35
Stability	0.41	0.46	0.51	0.53	0.53	0.48
Transfer to MOBY						
Interpolation to MOBY wavelengths	0.20	0.15	0.03	0.03	0.03	0.03
Reproducibility	0.37	0.39	0.42	0.44	0.42	0.30
Wavelength accuracy	0.29	0.08	0.04	0.03	0.01	0.04
Stray light	0.66	0.29	0.13	0.21	0.36	0.64
Temperature	0.25	0.25	0.25	0.25	0.25	0.25
MOBY stability during deployment						
System response	1.59	1.3	1.19	1.11	1.08	0.92
In-water internal calibration	0.43	0.42	0.44	0.46	0.51	0.55
Wavelength stability	0.13	0.14	1.12	0.82	1.37	0.65
Environmental						
Type A (good scans & all days)	4.1	4.4	4.5	4.4	4.0	3.2
Type A (good days only)*	*0.80*	*0.83*	*0.87*	*1.02*	*0.64*	*1.31*
Temporal overlap	0.3	0.3	0.3	0.3	0.3	0.3
Self-shading (uncorrected)	1	1	1.2	1.75	2.5	12
Self-shading (corrected)*	*0.2*	*0.2*	*0.24*	*0.35*	*0.5*	*2.4*
In-water bio-fouling	1	1	1	1	1	1
Combined Standard Uncertainty	4.8	4.9	5.1	5.1	5.2	12.6
Combined Standard Uncertainty*	*2.4*	*2.1*	*2.4*	*2.3*	*2.4*	*3.3*

Brown, et al., (2007) "The Marine Optical BuoY (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration." *Proc. SPIE* **6744**, 67441M1 – 67441M12.

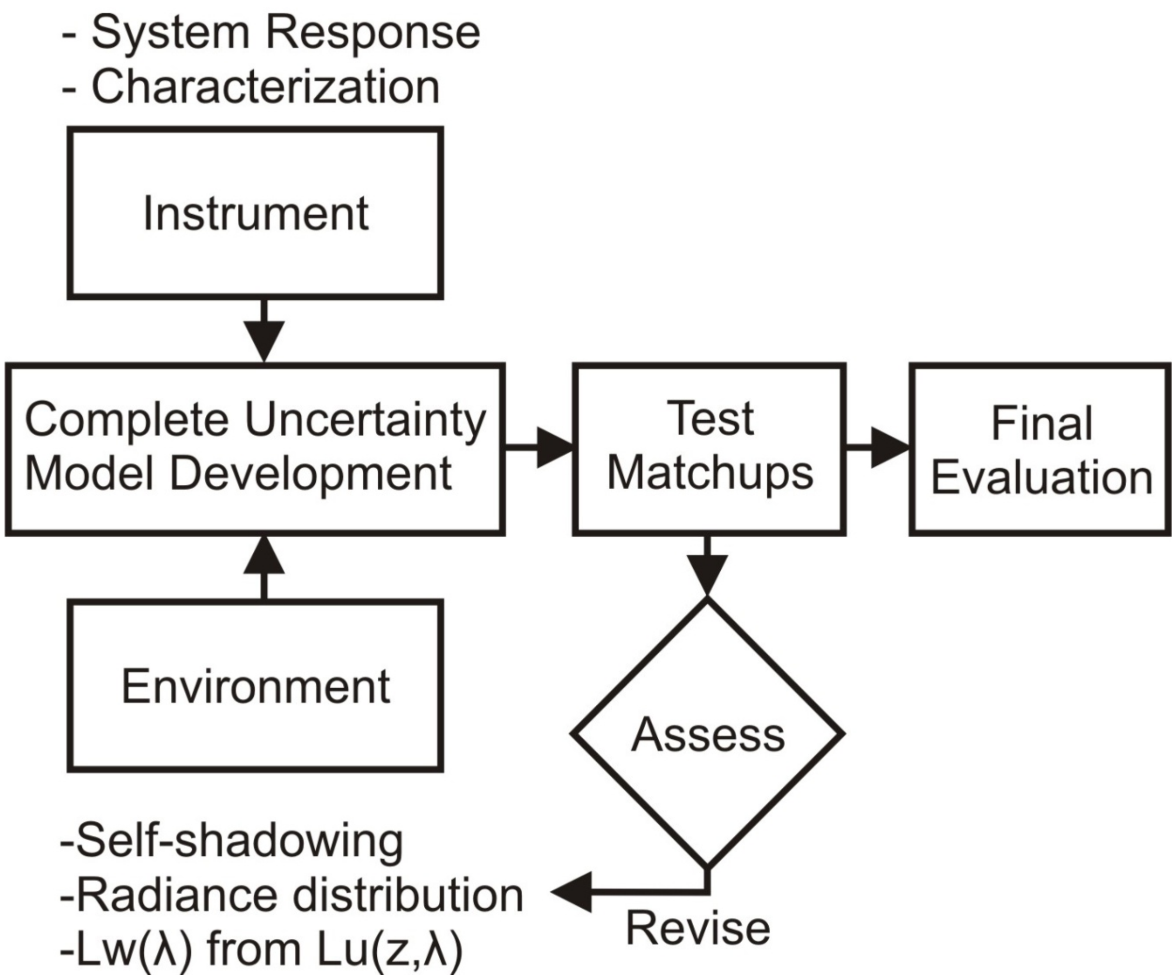
SORTIE In situ Intercomparison



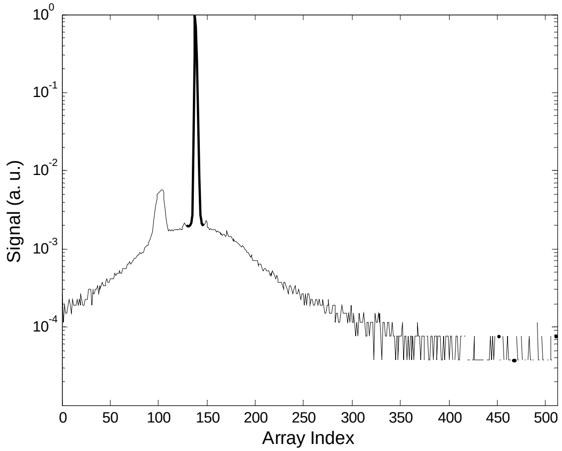


Voss et al. *JOAT*, **27**, 1747-1759 (2010).

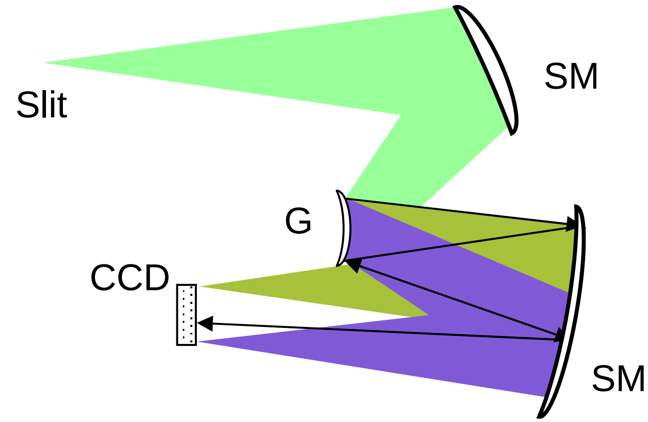
MOBY Uncertainty Analysis



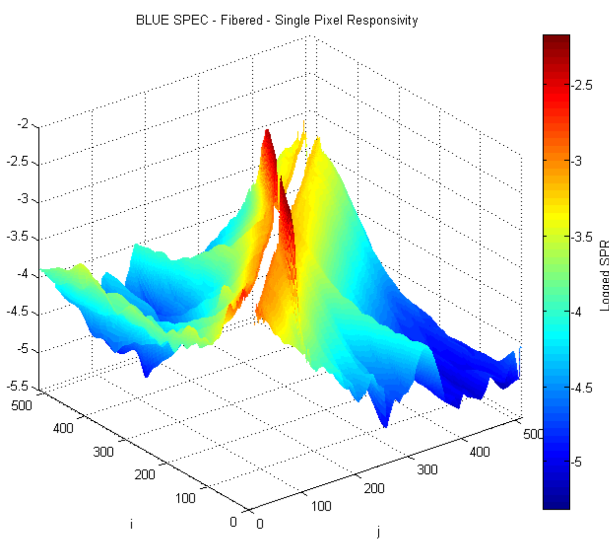
Stray Light



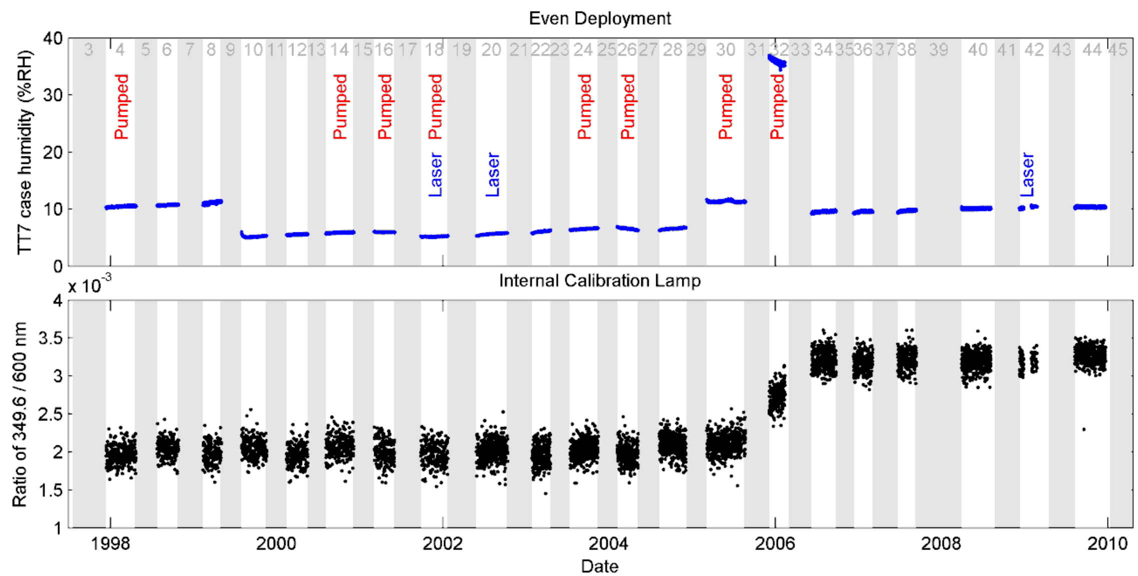
Signal (a. u.)
Array Index



Slit
SM
G
CCD
SM

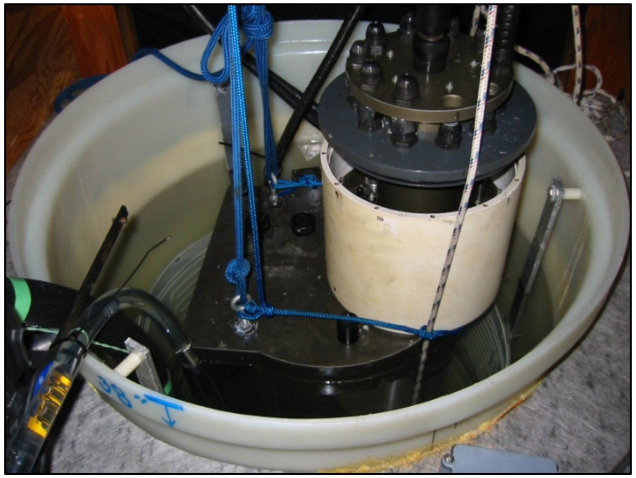


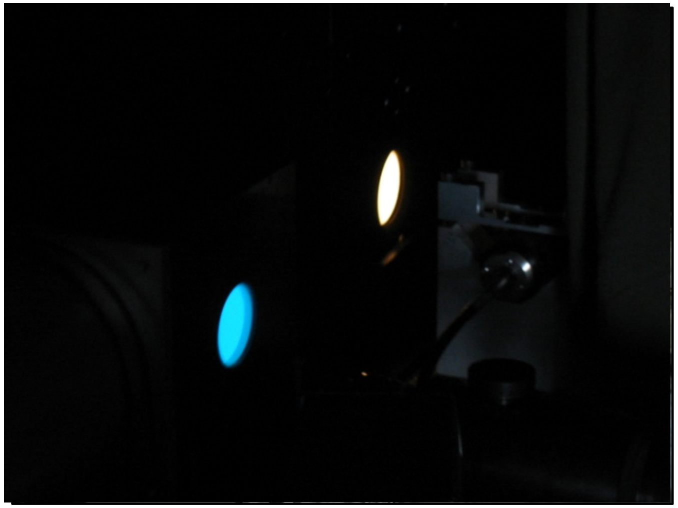
BLUE SPEC - Fibered - Single Pixel Responsivity
i
j

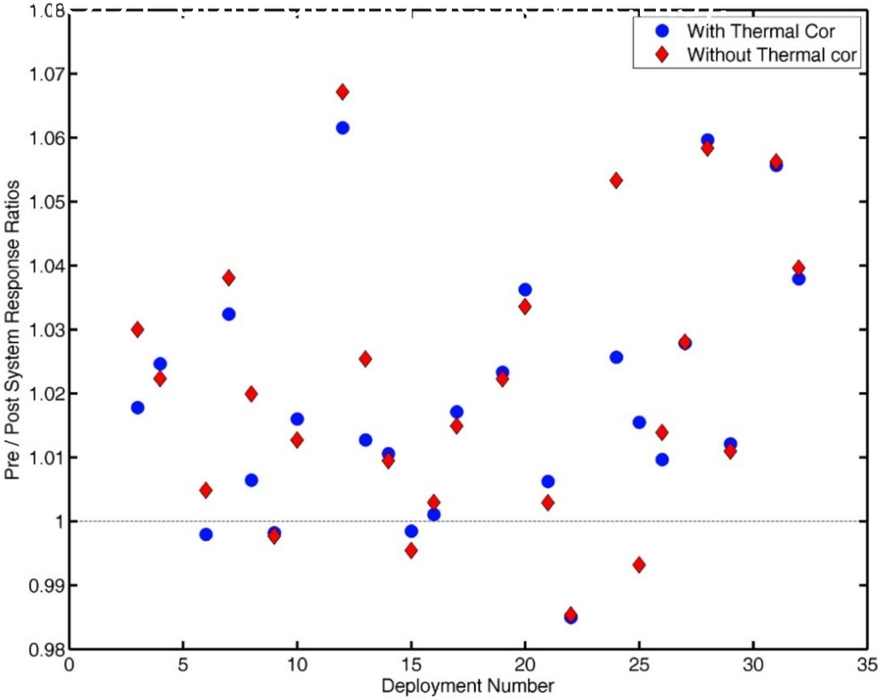


Even Deployment
TT7 case humidity (%RH)
Pumped
Laser
Internal Calibration Lamp
Ratio of 349.6 / 600 nm
Date

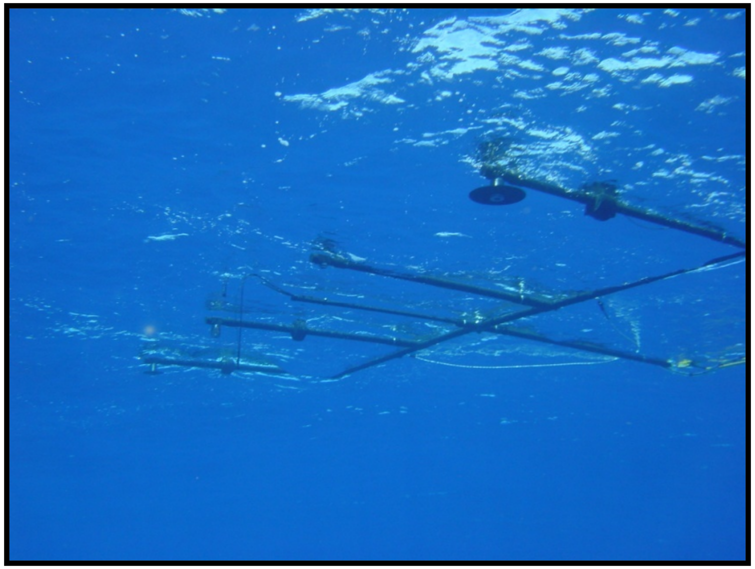
Temperature



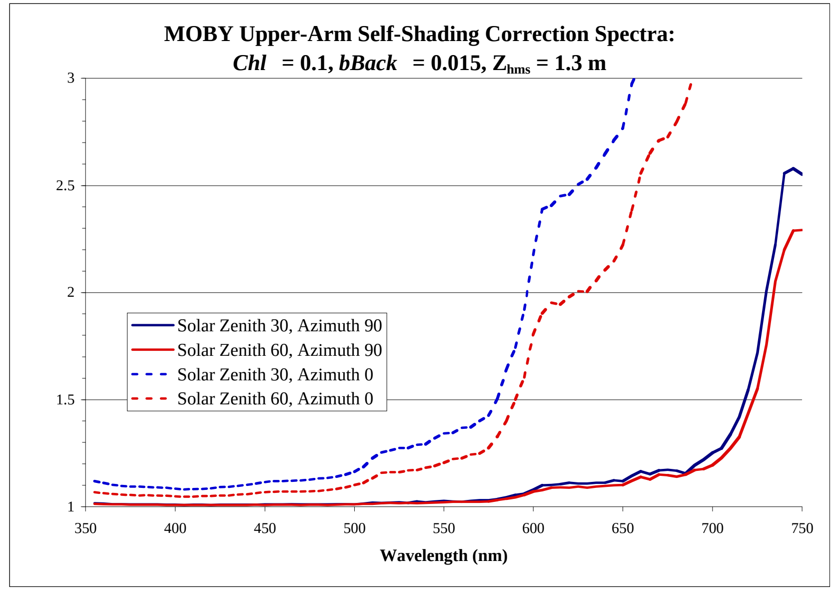


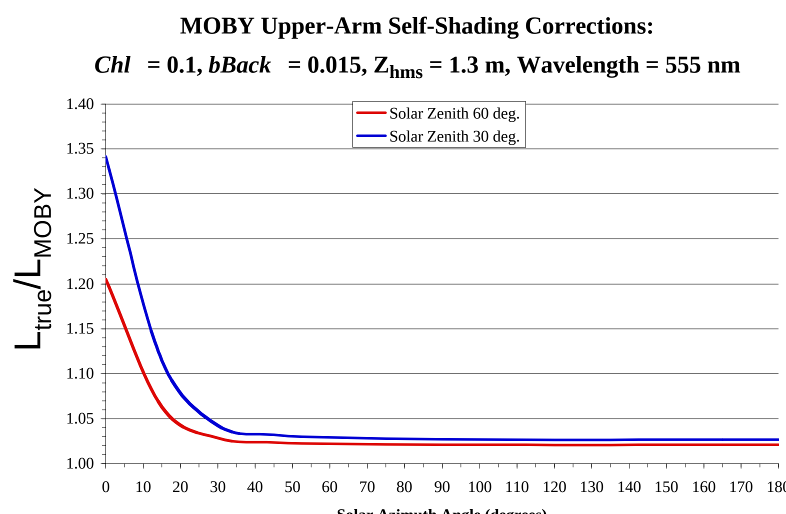


Self Shading



Field measurements using multi fiber “simultaneous” spectrograph



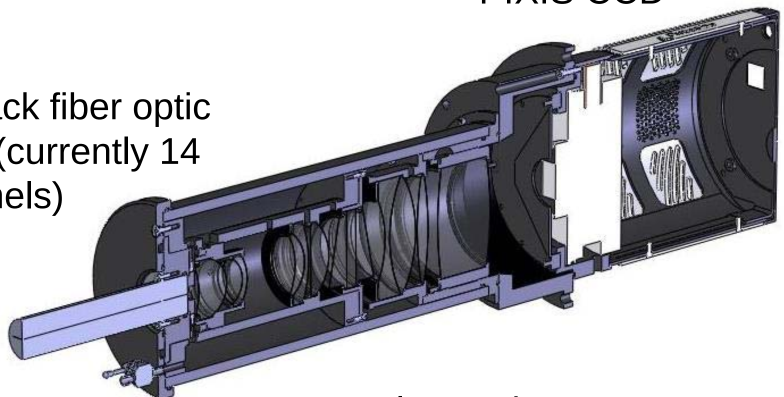


Mueller, J.L. (2007). Self-shading corrections for MOBY upwelling radiance measurements. Final Tech. Rpt. NOAA Grant NA04NES4400007, 33pp. (http://physoce.mlml.calstate.edu/moby/papers/papers/IssFinalRpt.pdf)

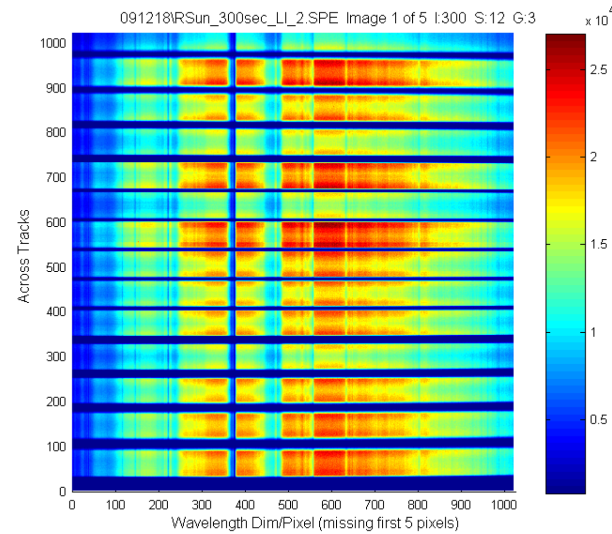
MOBY-C will be Simultaneous

Princeton Instrument
PIXIS CCD

Romack fiber optic input (currently 14 channels)

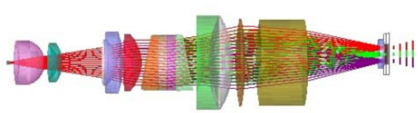


Resonon Volume Phase Holographic (VPH) in line spectrograph

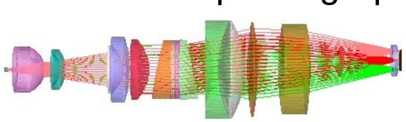


Es with red in line spectrograph

Blue In Line Spectrograph



Red In Line Spectrograph

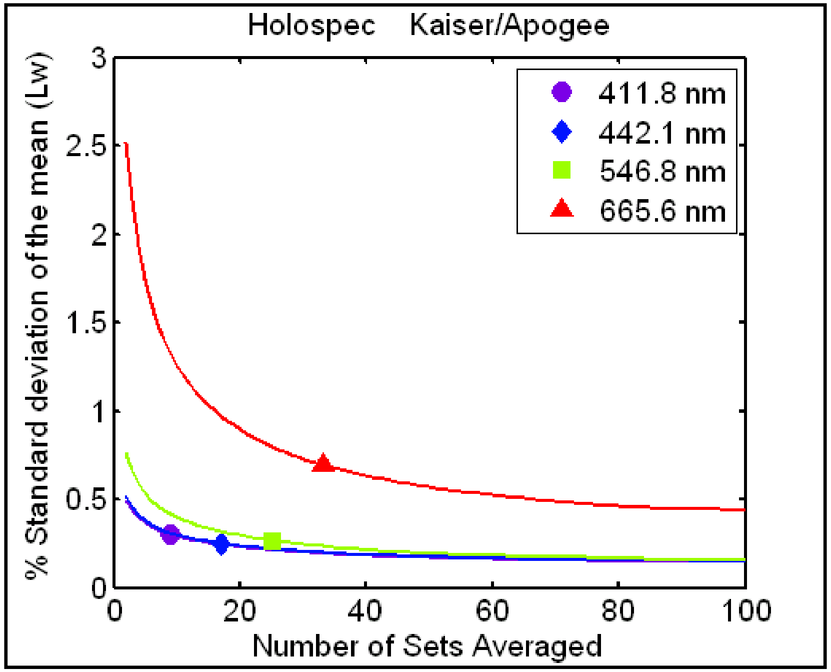


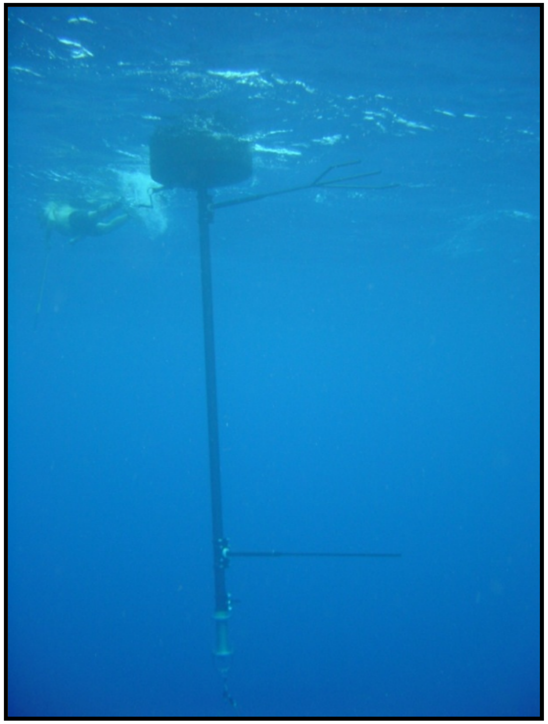
The inputs to the optical fibers sample different inputs (e.,g. depths). The fiber outputs are aligned vertically at the entrance slit. The prism-grating-prism in-line optical system (Resonon, Inc.) images the different input channels **at the same time** on the CCD camera), spaced along the slit direction.

The Type A uncertainty component will be reduced.

Increase in Sample Number – Type A

The simultaneous design places no restriction on the number of samples averaged. This is an advantage compared to MOBY, where the sequential measurements of Es, LuMid, Es, LuTop, Es, LuBot, Es, LuMOS, Es means an increase in number of samples would increase the time between these data collections, impacting the determination of Lw.



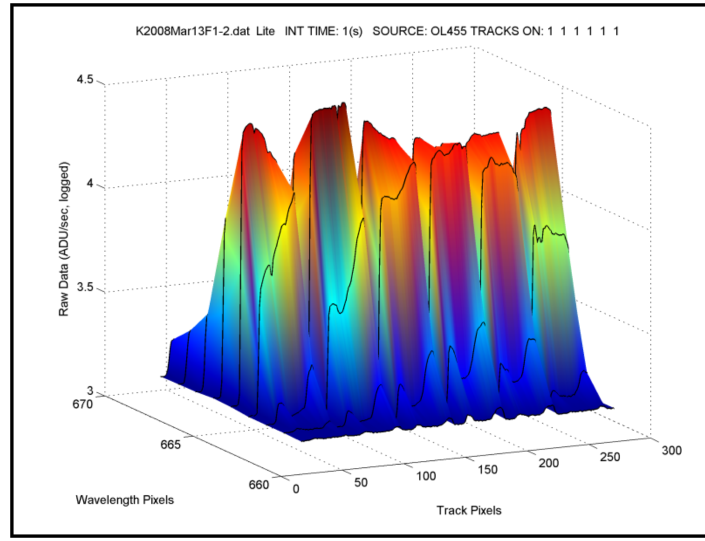


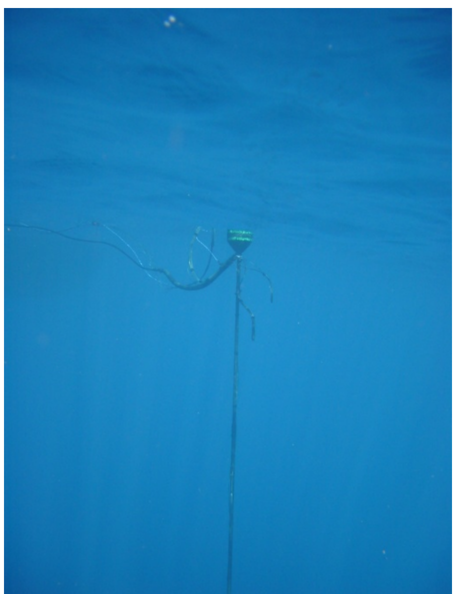
Band averaged results with a prototype 6-channel hyperspectral system tested in Case 1 waters off Oahu. Five to 100 scans were acquired with 4 sec integration times.

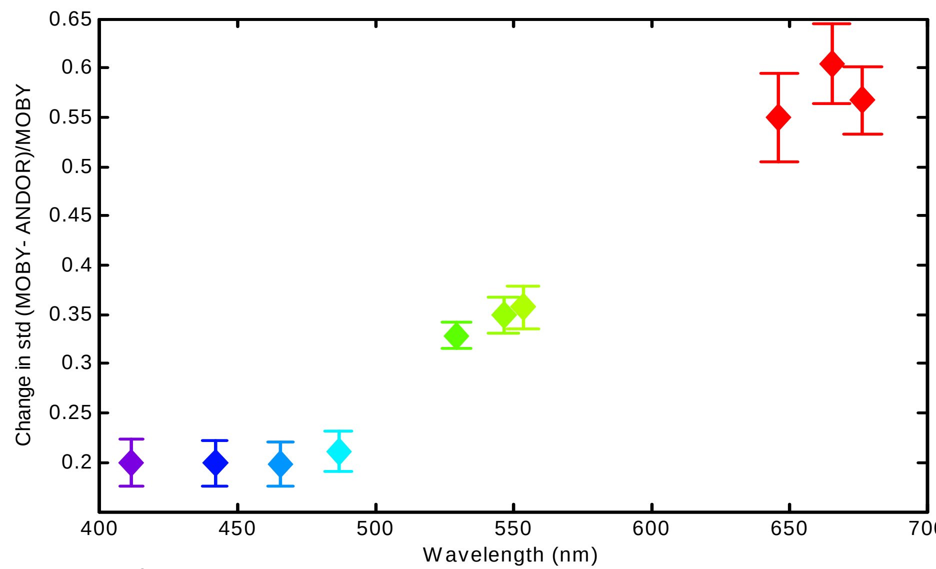
Yarbrough, M.A., S. Flora, M.E. Feinholz, T. Houlihan, Y.S. Kim, S.W. Brown, B.C. Johnson, K. Voss and D.K Clark (2007). "Simultaneous measurement of up-welling spectral radiance using a fiber-coupled CCD spectrograph." Proc. SPIE 6680, 66800J-1 to 66800J-11.

Effect of Correlated Variability – Type A

The effect of correlations in the light field was investigated by deriving $L_w(\lambda)$ from four simultaneous $L_u(\lambda)$s and by randomly sampling the $L_u(\lambda)$ scans in time to simulate the current MOBY sampling statistics. The measurement uncertainty was reduced between 20% to 60% for the ocean color bands.





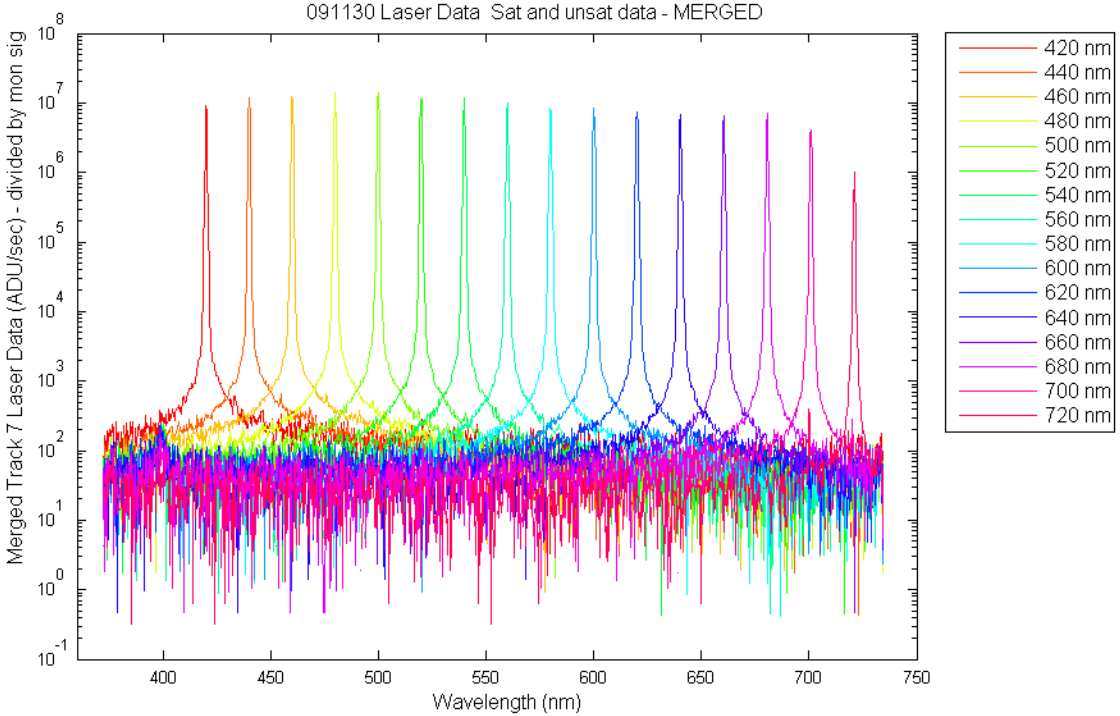


MOBY-C Spectrograph Parameters

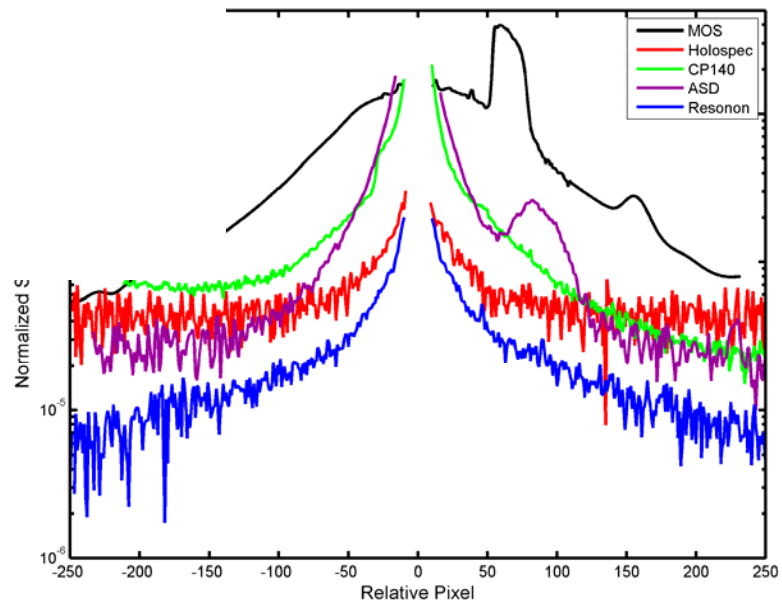
Parameter	Blue	Red
Size, cm	13.7 x 41.7	13.7 x 43.2
Spectral coverage, nm	370 - 720	500 - 900
Spectral resolution, nm	0.34	0.39
Image at focal plane, mm	13 x 13	13 x 13
Slit dimensions, mm	13 x 0.025	13 x 0.025
Thermal effect, pixel/deg C	< 0.05 pixel	< 0.05 pixel
MTF @ 38 line pr / mm	76 at 545 nm	61 at 700 nm
Throughput, %	74.8 at 430 nm	72.5 at 700 nm
Ghosting / Stray Light	< 0.5% at 420 nm	< 0.6% at 520 nm

"Report on Blue and Red Imaging Spectrometers for MOBY," Michael Kehoe and Casey Dodge, Resonon, Inc.

Laser Characterization on SIRCUS



Initial testing on SIRCUS; focus was on Track 7. Results are very encouraging, both in spectral and spatial dimension: ~10^{-5}.



The spectral stray light is the best we've seen so far for a single grating system. This is 20x better than MOS in terms of integrated area.

Thank You!