Executive Summary and SeaHARRE Results

Stanford Hooker
NASA/Goddard Space Flight Center
Greenbelt, Maryland

We Are Not Alone—*And Many Other Professionals Have a Much Tougher Job*

In the recertification of a hospital's ability to provide health care, the review commission investigates all departments and interviews as many key personnel as possible. Part of the process includes a reexamination of any so-called *sentinel events*, that is, those occurrences wherein the quality of patient care was significantly below established standards and everyone involved with the delivery of the relevant medical services needs to sit up and take notice. One of the first questions in such inquiries is

Did the patient die?

Whenever I try to explain an oceanographic problem that NASA is working on and how measurement accuracy and precision are being investigated to understand the source of the problem, someone maintaining quality assurance at a hospital usually asks this same question, "*Did anyone die?"* Obviously, the answer is, "*No*," and the rejoinder is almost always

You don't have a problem!

Calibration and Validation Activities Require a Quality Assurance (QA) Vigilance in all Disciplines

Large—*and completely avoidable*—uncertainties are a recurring part of calibration and validation activities, even for disciplines with extensive histories of trying to minimize them. The radiometric community, for example, has participated in absolute calibration and data processing round robins, but field protocols continue to be ignored by some practitioners.

The Problem Being Considered Involves Chromatography

Many complex problems defy simple and compact packaging, but in this case the basic difficulty involves expertise in chromatography, so a logical question is

What is the definition of a chromatographer?

Although there will be some variance in the answers to this question, it appears likely that most practitioners will agree on the following definition:

A chromatographer is capable of developing—and usually develops—an HPLC method.

With this as a definition, the people we need to listen to the most are John Dolan, Laurie Van Heukelem, and Crystal Thomas.

SeaHARRE Participants Summary

Extensive protocols and detailed uncertainty analyses were first established for radiometric measurements. No similar accomplishment exists for biogeochemical measurements except for HPLC, which was made possible by the contributions of the SeaHARRE community. The resulting round robins have emphasized international participation and the recruitment of new participants (green) and novice practitioners (yellow). There have also been specialized investigations of damaged samples (D), reanalyses to better understand uncertainties (R), and the use of two simultaneous methods for improved evaluation of methodological differences (2).

Code	Organization (and Country)	Principal Scientist	SH-1	SH-2	SH-3	SH-4
B	Bedford Institue of Oceanography (Canada)	Victoria Stuart				
C	Common. Scientific and Indust. Res. Org. (Australia)	Lesley Clementson				
D	DHI Water and Environment (Denmark)	Louise Schlüter				
F	USF/Florida Institute of Oceanography (USA)	Dave Millie				
G	NASA Goddard Space Flight Center (USA)	Mary Russ				
H	University of Maryland Horn Point Laboratory (USA)	Laurie Van Heukelem		D		
J	Joint Research Centre (Italy)	Jean-François Berthon			R	
L	Laboratoire d'Océanographie de Villefranche (France)	Hervé Claustre				
M	Marine and Coastal Management (South Africa)	Ray Barlow			R	
N	Dalhousie University (Canada)	Claire Normandeau				
P	Plymouth Marine Laboratory (United Kingdom)	Jim Aiken				
S	San Diego State University/CHORS (USA)	Charles Trees			2	
U	University of South Carolina (USA)	Jay Pinckney				

SeaHARRE Method Diversity as a Function of Time (and for a TChl *a* range of 0.02 – 42.7 mg m^{-3})

The SeaHARRE field sampling has emphasized a large variety of ecosystems (oligotrophic gyres, wind driven upwelling, etc.), but most of them have been in open ocean (Case-1) waters. In addition, a diversity of methods have been used by SeaHARRE participants, but the majority of them have been based on C_8 columns. Consequently, there has been a recurring emphasis to add new practitioners who are using C_{18} methods, so method diversity can be maintained over time.

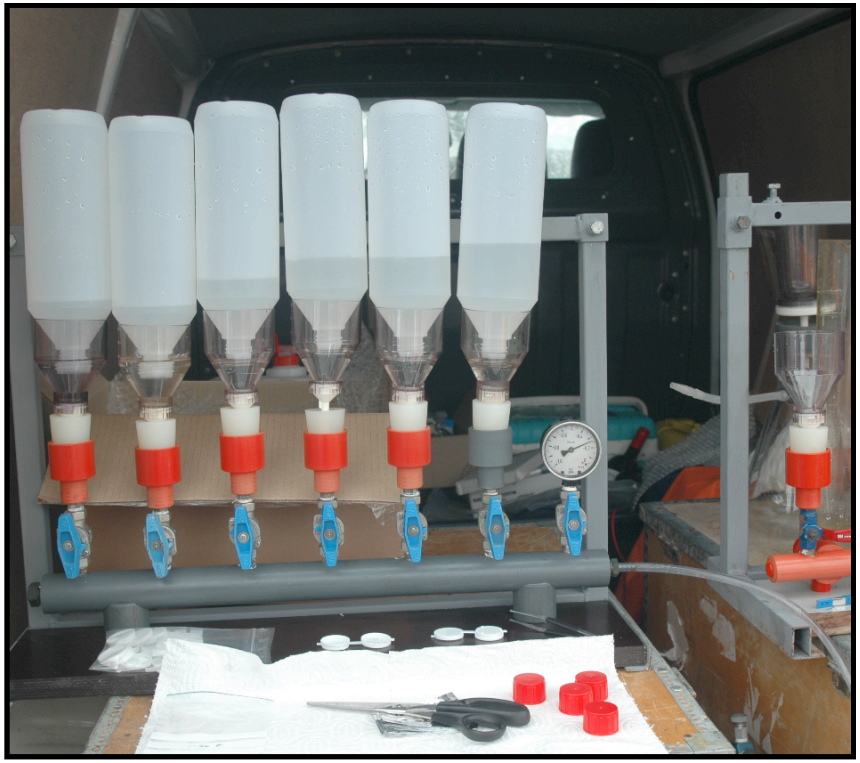
Column Type	Method Citation	Open Ocean (Case-1) SH-1	SH-2	SH-3	Coastal SH-4	Total Labs
C_{18}	Gieskes and Kraay (1989)		B			1
C_{18}	Wright et al. (1991)	J	C,D,S	J,S18	N,S	8
C_{18}	Pinckney et al. (1996)				F,U	2
C_8	Vidussi et al. (1996)	L	L			2
C_8	Barlow et al. (1997)	M	M,P	M		4
C_8	Van Heukelem and Thomas (2001)	H	H	C,D,H,L,S8	C,D,G,H,J,L	13

The laboratory codes for new participants are shown in red.

Despite an explicit effort to maximize method diversity and a strong initial desire to not have the SeaHARRE community produce a unified method, there has been a significant movement by the analysts to adopt one method — *Van Heukelem and Thomas (2001)* — more than any other. Consequently, a majority of the SeaHARRE data comes from the use of a C_8 method. This evolution was the reason for soliciting as many new practitioners using a C_{18} method as possible.

Field Sampling for SeaHARRE-4—*The First Activity in Coastal (Case-2) Waters*

The emphasis for SeaHARRE-4 was on coastal (Case-2) waters. The sample set includes 12 different locations from the fjords, estuaries, and bays of Denmark. All samples were collected in triplicate and distributed in November 2006.



The sampling plan included a concerted effort to obtain the widest range in water properties possible (8 – 28 PSU) plus a diversity of phytoplankton populations and sizes (including blooms dominated by a single species) to ensure the most complex mix of pigments possible. *At some level, no one area is sufficient, but at another level, any one area is typical as long as the range in complexity of the coastal environment is captured.*

SeaHARRE-4 Participants and Analysis

The laboratories represented in SeaHARRE-4 are a mixture of established and new HPLC practitioners as well as established and new round-robin participants. The new additions have well-established expertise in coastal sampling. Every effort was made to increase the diversity of international groups (e.g., a concerted effort was made to include a South American institute) and methods (e.g., the Zapata method), but the timing of the activity was not necessarily advantageous to the invitees. All of the participants agreed to make an additional analysis with the HPLC extracts to ensure a more comprehensive use of the samples.

Sample Set	*Institute or Laboratory*	*Principal Scientist*	*Country*	*Lab. Code*	*HPLC Pigs*	*Fluor. Chl a*	*Spec. Chl a*	*Absorp-tion*	*Method*
1	CSIRO	L. Clementson	Australia	C	H		S	A	Van Heukelem and Thomas
2	DHI	L. Schlüter	Denmark	D	H	F	S		Van Heukelem and Thomas
3	GSFC/UMBC	M. Russ	USA	G	H		S		Van Heukelem and Thomas
4	HPL	L. Van Heukelem	USA	H	H		S		Van Heukelem and Thomas
5	HPL			H'	H	F		A	Van Heukelem and Thomas
6	JRC	J-F. Berthon	Italy	J	H		S		Van Heukelem and Thomas
7	LOV	H. Claustre	France	L	H	F	S		Van Heukelem and Thomas
8	LOV			L'	H	F	S	A	Van Heukelem and Thomas
9	USC	J. Pinckney	USA	U	H	F	S		Pinckney et al.
10	USF/FIO	D. Millie	USA	F	H	F	S		Pinckney et al.
11	SDSU/CHORS	C. Trees	USA	S	H	F			Wright et al.
12	Dalhousie Univ.	C. Normandeau	Canada	N	H	F			Wright et al.
12	*10*	*10*	*6*	*12*	*12*	*8*	*9*	*3*	*4*

Establishing the QA Subset to Ensure a Proper Referencing System for Computing Uncertainties

The first step in the analysis of the SeaHARRE data is to establish the QA subset. This is initially based on the precision obtained with the field samples. A laboratory with an average precision not satisfying *semiquantitative* analysis—more than 8% plus 2% for field sample variability—is excluded from the QA subset (F and U). In addition, a laboratory with three or more primary pigments with a precision exceeding routine capabilities (13%) is considered for exclusion (N). For the first time, all the C_{18} methods failed to qualify for the QA subset, and the exclusion of CHORS (S) is considered in the final test (which also eliminated J).

The precision values shown in yellow highlights are artificially low, the entries shown in red exceed the 10% threshold, and the laboratories shown in blue highlights are the QA subset (denoted A').

Lab.	TChl *a*	TChl *b*	TChl *c*	Caro	But	Hex	Allo	Diad	Diato	Fuco	Peri	Zea	Ave.
C	4.3	2.9	3.6	2.8	1.4	1.9	4.2	4.3	8.9	4.6	4.4	4.2	4.0
D	3.6	8.2	4.8	6.0	0.5	5.1	4.9	5.6	12.2	5.8	4.7	5.9	5.6
F	8.6	12.6	61.9	16.6	35.5	39.8	33.5	12.6	28.0	14.3	34.0	26.4	27.0
G	4.8	5.9	3.4	4.9	0.0	0.1	5.7	5.1	4.1	4.3	3.7	6.1	4.0
H	5.8	4.9	5.5	5.9	17.9	4.1	5.4	7.3	8.0	5.0	8.3	4.4	6.9
J	3.5	6.8	11.1	6.1	6.1	4.3	4.7	5.8	22.7	4.0	3.5	6.4	7.1
L	2.8	5.0	4.3	2.8	2.9	2.0	10.9	4.0	7.2	3.6	3.3	5.1	4.5
L'	1.8	5.7	2.4	3.1	6.2	1.9	4.5	8.9	5.9	2.6	2.7	4.1	4.2
N	7.5	6.6	9.4	18.6	1.0	3.8	16.3	16.7	9.2	10.2	6.4	5.7	9.3
S	2.6	3.5	6.1	4.1	3.3	6.5	4.9	5.9	14.0	2.5	8.8	5.9	5.7
U	6.9	19.6	12.2	9.5	0.0	0.2	15.2	9.0	23.2	8.9	5.9	26.2	11.4
A Ave.	4.7	7.4	11.3	7.3	6.8	6.4	10.0	7.8	13.0	6.0	7.8	9.1	8.1
A' Ave.	3.9	5.5	4.0	4.3	7.1	3.0	5.9	5.9	7.7	4.3	4.5	5.0	5.1
A^+ Ave.	5.8	9.8	20.1	11.0	9.2	10.9	14.9	10.0	19.4	8.0	11.7	14.1	12.1

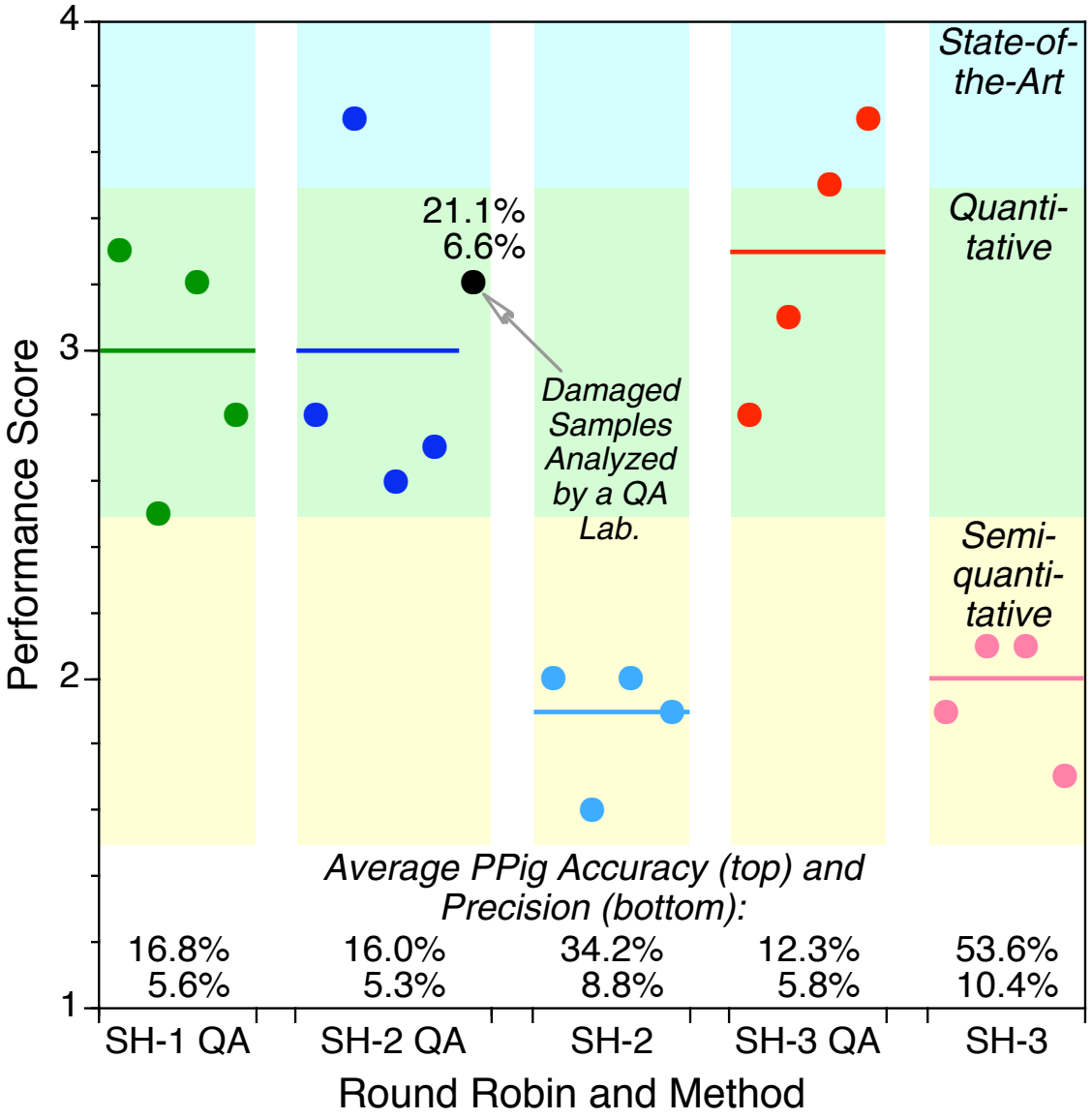
Final Test for Establishing the QA Subset to Ensure a Proper Referencing System

The laboratories that are part of the initial QA subset are then tested individually to see if the average uncertainties for each are within the performance requirements for *semiquantitative* analysis (15% and 25% accuracy for TChl *a* and the primary pigments, respectively). In addition, if three or more pigments have numerous uncertainties exceeding 100%, a laboratory is considered for exclusion (depending on the magnitude of the large uncertainties and the performance of the other laboratories). Although CHORS (S) had excellent precision, the accuracy of the results were not in keeping with semiquantitative analysis, even if the most problematic pigments (But and Hex) are ignored.

Sample	TChl *a*	TChl *b*	TChl *c*	Caro	But	Hex	Allo	Diad	Diato	Fuco	Peri	Zea	Ave.	NoB&H
B	20	18	36	162	100	632	22	24	83	9	384	30	127	79
C	36	13	51	72	98	250	7	14	53	14	178	17	67	46
D	59	24	31	74	17913	8182	1	19	3	30	2608	187	2427	303
E	27	31	25	68	93	12518	9	20	19	4	90	102	1084	40
F	23	24	52	58	48	9838	2	11	19	1	21	37	845	25
G	26	28	30	53	74	2820	4	4	4	1	28	89	263	27
H	38	25	27	64	441	150	2	25	24	14	14	25	71	26
I	44	37	35	57	9372	70	4	12	15	10	14	20	808	25
J	55	23	61	78	1455	185	3	5	54	2	18	98	170	40
K	29	28	18	45	1996	1639	1	4	24	2	20	47	321	22
L	21	27	14	36	1283	524	5	11	67	2	18	4	168	20
M	30	10	8	42	127	2	2	14	2	1	17	6	22	13
S Ave.	34	24	32	67	2750	3068	5	14	31	8	284	55	531	55
J Ave.	23	21	13	10	442	66	32	18	33	17	58	30	64	25

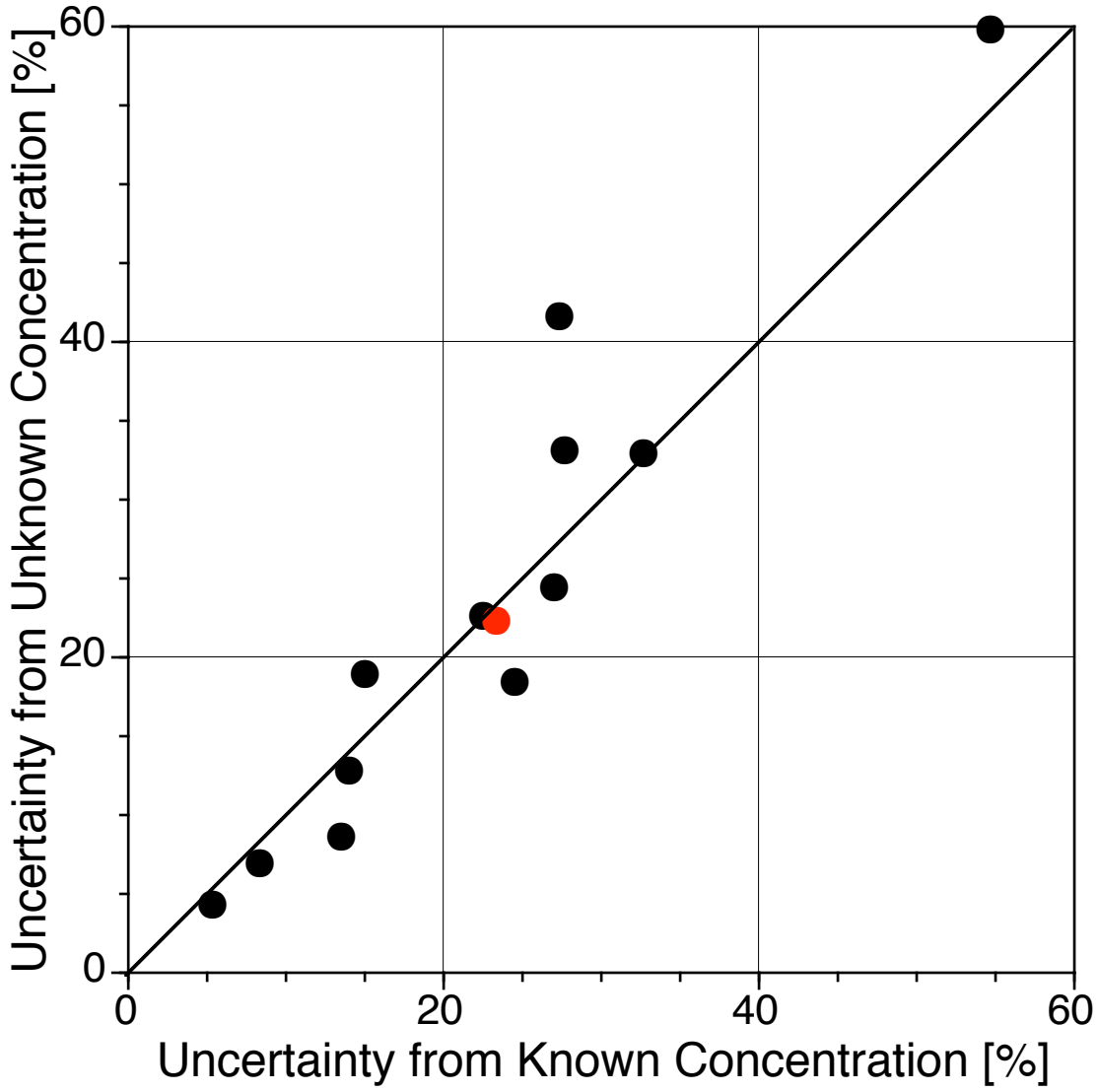
Validation of Establishing the QA Subset from the Performance Metrics of the Individual Laboratories

Once the analyses for a round robin are completed, the performance metrics are compared to what the investigators first submitted and adjusted according to what was revealed during the activity. The composition of the QA subset is rechecked to make sure all the members within the QA subset performed at the quantitative level or higher. For the round robins executed so far, the determination of the QA subset has always been found to be in compliance the final performance metrics. *Even when a QA lab analyzed a damaged set of samples.*



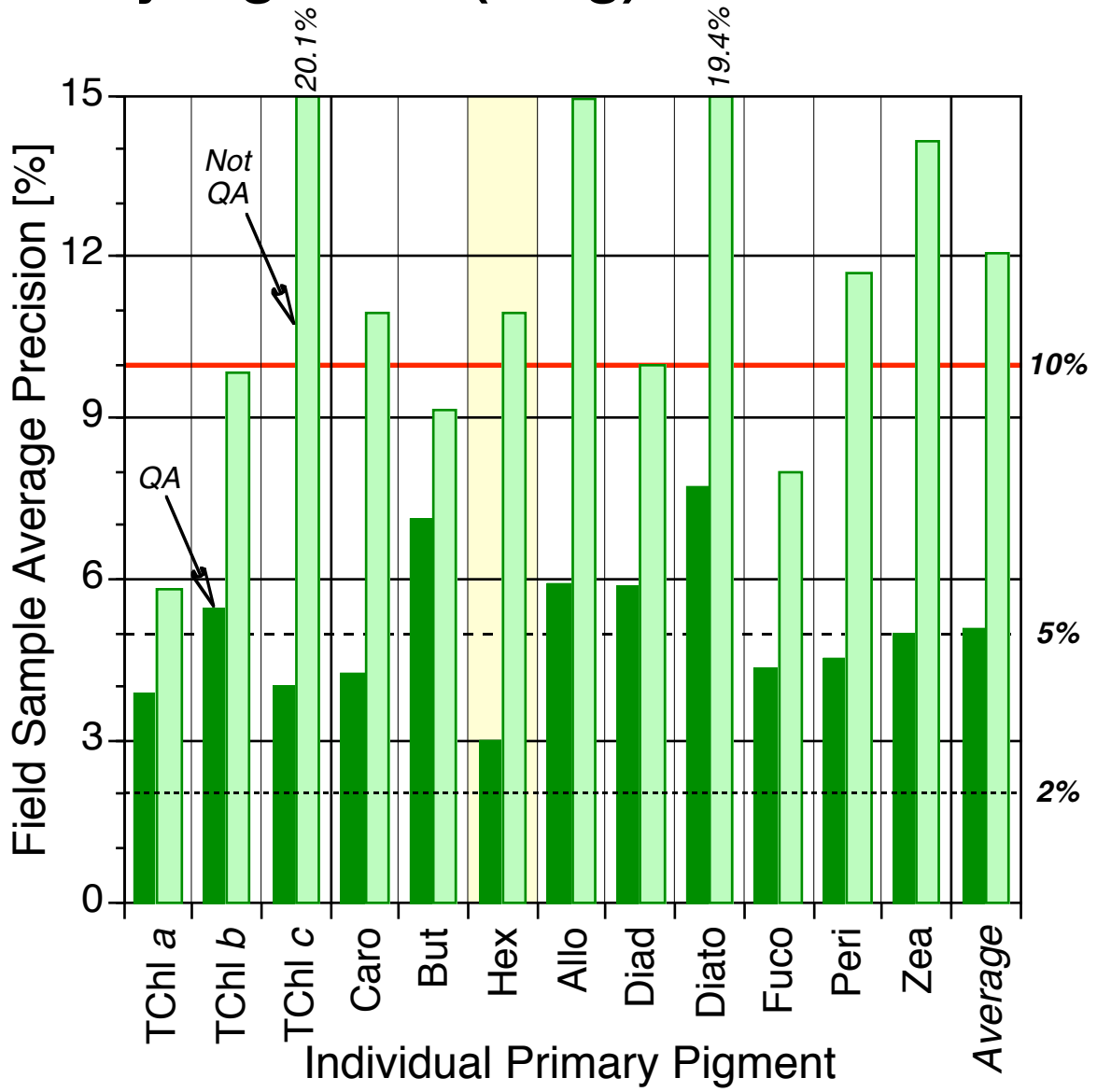
Validation of the Referencing Scheme Using the Results from Mixed Standards

The validity of the referencing system used to compute accuracy (and, thus, uncertainties) can be investigated using mixed standards (a single solution containing a variety of standards all mixed together in known concentrations). The validation occurs by comparing the uncertainties in the pigment concentrations from the various methods computed using a) the known concentrations within the mix, and b) the average pigment concentrations derived from all the methods. The average uncertainties from these two approaches (the red dot in the figure) agree to within 1.5%.



SeaHARRE-4 Average Precision for the Individual Primary Pigments (PPig)

The SeaHARRE-4 PPig precision values exhibit a strong dependence on whether or not the results are part of the QA subset. Except for TChl *a* and Fuco, all of the results from the laboratories not in the QA subset have average precisions very close to or in excess of 10%. The precision for Hex in the QA subset is anomalously low, because an unusually large number of the results were at the level of no detection (which adds a very large number of zero precision values to the averages).



Details of the PPig Uncertainties in the QA Subset and Compared to All SeaHARRE Activities

Sample	TChl *a*	TChl *b*	TChl *c*	Caro	But	Hex	Allo	Diad	Diato	Fuco	Peri	Zea
B	3	11	33	2	123	116	54	13	11	6	149	6
C	12	12	32	3	117	129	11	9	46	7	112	11
D	4	16	22	9	0	0	8	6	37	6	125	71
E	4	10	19	6	111	0	9	7	35	5	114	5
F	6	8	16	7	72	0	6	19	65	5	113	7
G	3	8	15	9	111	97	9	8	34	6	94	4
H	4	17	13	6	35	45	8	4	15	4	10	15
I	8	30	20	12	98	34	20	16	23	9	10	16
J	12	10	19	22	123	81	12	9	102	8	15	64
K	6	5	18	9	84	92	7	7	101	8	9	6
L	6	7	13	9	151	67	6	8	62	5	10	15
M	11	34	12	10	31	5	6	10	31	7	17	12
A' Ave.	7	14	19	9	88	56	13	10	47	6	65	19
SH-1 A'	7	14	26	18	24	25	39	16	56	9	13	11
SH-2 A'	6	16	22	17	31	10	20	9	21	5	15	21
SH-3 A'	6	14	15	13	15	6	4	5	18	11	30	10

The details of the PPig uncertainties in the QA subset show six pigments were quantitated at a high level of accuracy and four were not (But, Hex, Diato, and Hex)—the remaining two pigments are at levels consistent with prior round robins. The good results for the former are mostly expected (particularly TChl a and Fuco), but the poor results for the latter are much worse than has been seen in the past (except Diato during SeaHARRE-1). *The steady improvements achieved in Case-1 waters (SH-1 through SH-3) have now been significantly challenged*.

Details of the SeaHARRE-4 PPig Uncertainties for the Laboratories Not in the QA Subset

Sample	TChl *a*	TChl *b*	TChl *c*	Caro	But	Hex	Allo	Diad	Diato	Fuco	Peri	Zea
B	13	20	477	257	111	243	61	34	69	36	161	38
C	20	36	375	674	89	118	39	21	60	21	196	35
D	24	15	257	304	91876	2434	24	10	51	21	722	569
E	19	14	359	224	98	2938	20	32	58	23	83	56
F	14	16	291	132	444	2766	13	10	83	15	71	41
G	14	15	245	250	789	662	10	16	47	17	70	61
H	20	19	185	350	146	82	23	19	46	10	13	38
I	25	18	204	521	1996	53	24	16	27	11	11	27
J	34	18	214	365	415	243	16	27	75	33	12	111
K	18	15	230	338	579	415	20	10	53	13	7	27
L	16	15	197	190	325	158	14	21	147	13	14	19
M	22	34	161	465	71	8	18	17	41	15	59	33
A^+ Ave.	20	20	266	339	8078	843	23	19	63	19	118	88

The details of the PPig uncertainties for the laboratories not in the QA subset involve uncertainties that are much higher than has been seen in the past, wherein poor results were on the order of a small number of hundreds of percent. Results in the many hundreds of percent did occur, but were isolated; and results in the thousands of percent were even more sporadic. Very large uncertainties are almost always associated with a significant over-quantitation of a pigment. The likely culprits are usually coelution problems or a complete misidentification of a peak. Calibration mistakes can produce overestimation, but they cannot explain intermittent and very high uncertainties like those seen in But or Hex.

Summaries of the SeaHARRE-4 PPig Uncertainties—*Including a Novice Practitioner, G*

Lab.	TChl *a*	TChl *b*	TChl *c*	Caro	But	Hex	Allo	Diad	Diato	Fuco	Peri	Zea	PPig
C	4	16	31	7	79	78	27	16	56	12	53	21	33
D	5	14	30	18	65	46	9	8	31	6	104	17	30
G	8	7	9	4	76	62	14	8	66	3	52	21	28
H	7	7	17	12	145	57	9	7	64	8	68	15	35
L	7	18	15	6	71	45	6	7	32	3	57	20	24
L'	8	23	14	4	93	46	12	12	32	5	57	21	27
A' Ave.	7	14	19	9	88	56	13	10	47	6	65	19	29
F	25	20	55	1318	36471	123	30	18	87	33	93	57	3194
J	23	21	13	10	442	66	32	18	33	17	58	30	64
N	7	26	12	9	653	897	20	21	94	16	95	271	177
S	34	24	32	67	2750	3068	5	14	31	8	284	55	531
U	11	8	1218	293	76	63	30	25	72	21	61	25	159
A^+ Ave.	20	20	266	339	8078	843	23	19	63	19	118	88	825

The details of the PPig uncertainties for all the laboratories show many of the pigments for the laboratories not in the QA subset are within the requirements for calibration and validation activities (blue): 15% for TChl *a* and 25% for all the other pigments. Similarly, many of the results for the QA subset exceed the thresholds for calibration and validation (red), and the PPig average is above 25%. The latter is a direct consequence of the difficulties in quantitating But, Hex, Diato, and Peri; these same pigments are also at persistently high uncertainty levels in the laboratories that are not part of the QA subset (A^+). The A^+ subset is further degraded by the individual difficulties with TChl *c* (U) and Caro (F).

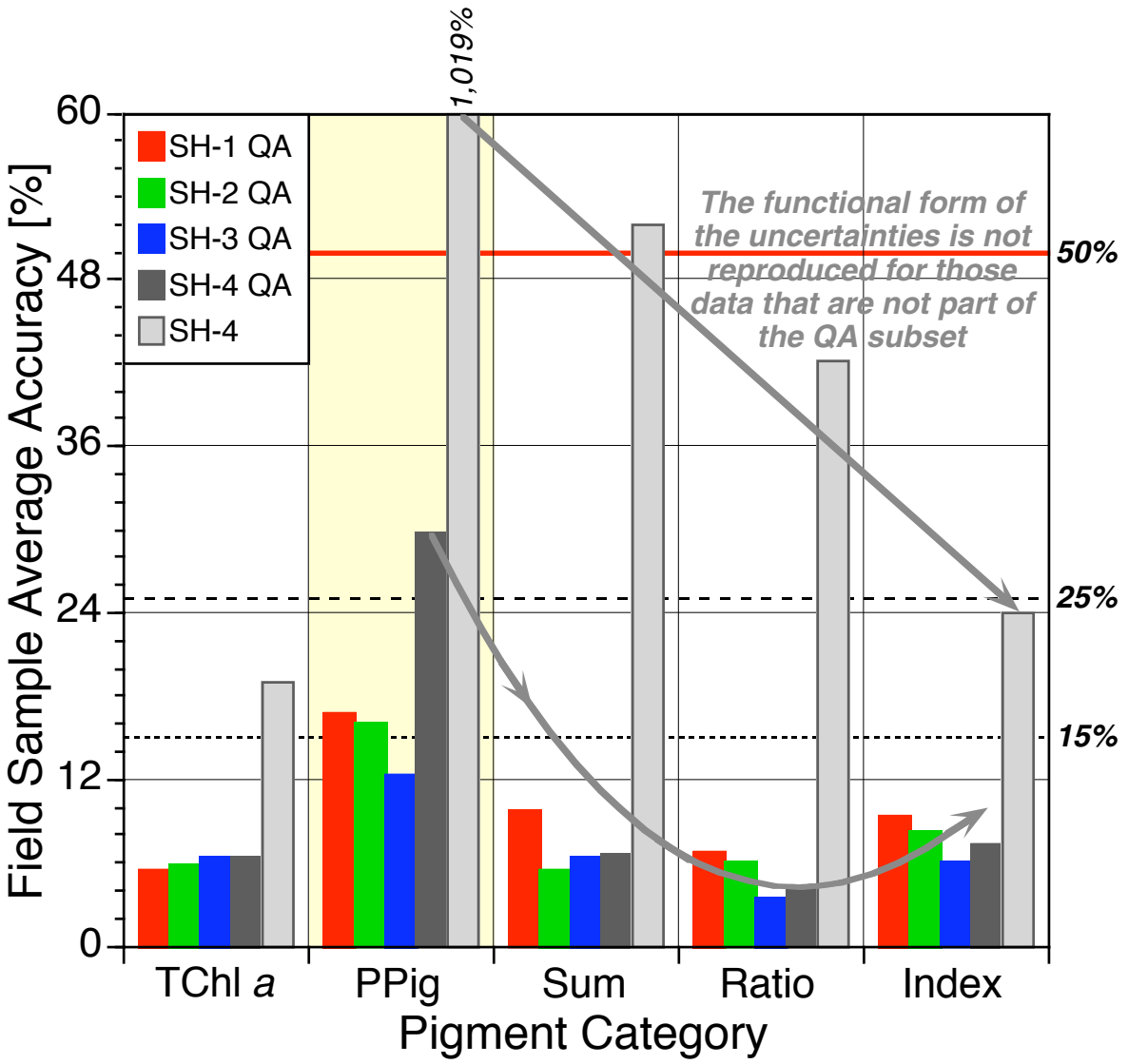
Summaries of the SeaHARRE-4 PPig Uncertainties for the DHI Mix

Lab.	TChl *a*	TChl *b*	TChl *c*	Caro	But	Hex	Allo	Diad	Diato	Fuco	Peri	Zea	PPig
C	6	6	24	5	3	6	5	1	1	2	1	1	5
D	6	2	8	6	8	6	1	8	2	3	3	17	6
G	3	5	6	4	2	1	5	3	2	2	11	5	4
H	2	10	5	0	7	3	4	5	4	4	7	9	5
L	4	8	5	3	0	2	2	1	5	5	8	3	4
A' Ave.	4	6	10	4	4	4	3	3	3	3	6	7	5
F	5	14	39	996	30	26	7	17	29	28	15	20	102
J	13	26	4	10	2	9	14	1	25	3	8	0	10
N	12	6	20	29	6	20	13	22	1	3	21	100	21
S	6	5	5	12	8	10	13	2	6	4	17	8	8
U	5	16	639	154	7	2	5	3	6	1	12	10	72
A^+ Ave.	8	13	142	240	11	13	11	9	13	8	15	27	42

Another aspect of SeaHARRE intercomparisons is the use of a mix of pigments designed to simulate a natural sample, but at concentrations significantly above the limit of detection of any method. In fact, the signal-to-noise ratio for the pigments is so high, the uncertainty in computing the concentrations in the mix is not driven by differences in how analysts interpret a chromatogram—the uncertainty is mostly due to calibration and coelution problems. Consequently, the uncertainty for each pigment should be to within 10%, and a laboratory performing within the QA subset should have an average PPig uncertainty of approximately 5%. The A^+ subset is distinguished here by uncertainties above 10% for the majority of the pigments.

SeaHARRE-4 Average Accuracy for the Higher-Order Variables

The accuracy of the higher-order variables (sums, ratios, and indices) for the QA subset are very similar to the other three round robins, even though accuracy for the PPig pigments does not meet the calibration and validation requirement. The reason for this result is the pigments responsible for degrading accuracy are at rather low concentrations, so they do not influence the higher-order products very much. For the results not in the QA subset, the notable results are the very high uncertainties and the absence of the functional form in the uncertainties.



Confirmation of Laboratory Performance Within the Tertiary Pigments

Sample	Lut	Neo	Phytin *a*	Phide *a*	Pras	Viola	Sample	Lut	Neo	Phytin *a*	Phide *a*	Pras	Viola
B	2	24	36	32	159	14	B	84	109	25	74	108	55
C	79	140	51	7	0	50	C	269	1058	17	35	9257	92
D	73	98	48	27	108	12	D	222	362	37	60	269	34
E	73	9	53	48	5	16	E	175	42	19	86	36	35
F	9	11	50	60	9	8	F	62	37	47	85	38	45
G	16	4	42	55	4	11	G	88	42	15	80	22	24
H	5	15	56	46	80	6	H	55	68	31	88	62	42
I	6	24	49	56	37	17	I	50	83	45	77	48	55
J	59	85	23	85	122	10	J	115	108	41	142	68	69
K	64	13	71	133	5	10	K	542	52	120	110	33	43
L	27	19	39	71	36	8	L	115	25	51	80	25	31
M	37	28	37	47	36	24	M	61	70	58	64	58	51
A' Ave.	37	39	46	56	50	16	A^+ Ave.	153	171	42	82	835	48

The tertiary pigments are usually characterized by poorer accuracy. For the QA subset the average accuracy is 40.7% and Viola is the only pigment satisfying the 25% uncertainty threshold. For the laboratories not in the QA subset, the average accuracy is 222.0% and no pigments satisfy the 25% requirement. For both types of data, there are significant indications of false positives (uncertainties above 100%) and false negatives (uncertainties just below 100%), although both effects are more pronounced for the laboratories not in the QA subset.