The role of *in situ* bio-optical data in ocean color satellite climate data record development

P. Jeremy Werdell, Sean W. Bailey, and Charles R. McClain
NASA Goddard Space Flight Center http://oceancolor.gsfc.nasa.gov jeremy.werdell@gsfc.nasa.gov

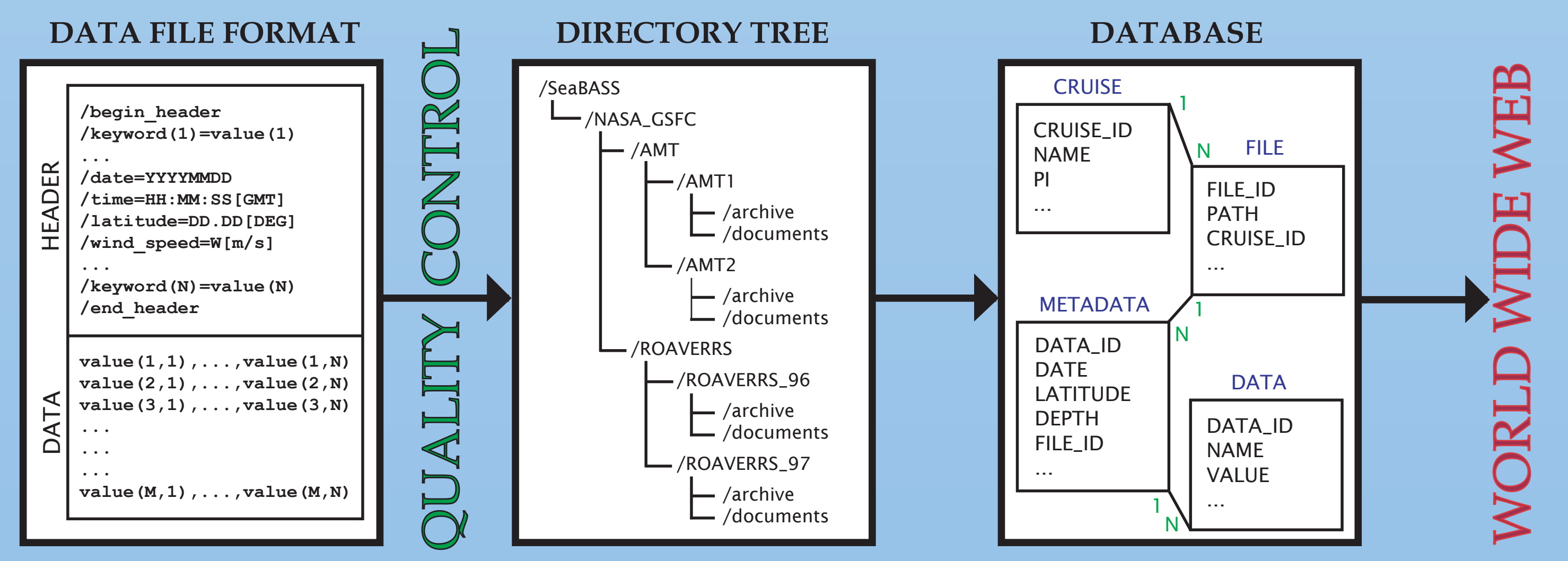
OS35B-05

1. INTRODUCTION

High quality *in situ* observations are prerequisite for satellite data product validation, remote-sensing algorithm development, and climate research on scales ranging from regional ecosystems to decadal global change. The NASA Ocean Biology Processing Group (OBPG) maintains a local repository of *in situ* marine and atmospheric bio-optical data, the SeaWiFS Bio-optical Archive and Storage System (SeaBASS; Werdell and Bailey 2002) to facilitate ocean color satellite validation efforts. Recently, the OBPG reevaluated these data and developed a publicly available data set for bio-optical algorithm development, the NASA bio-Optical Marine Algorithm Data set (NOMAD; Werdell and Bailey 2005). NOMAD includes over 3,400 stations of coincident radiometric spectra and phytoplankton pigment concentrations, encompassing chlorophyll *a* concentrations, C_a, ranging from 0.012 to 72.1 mg m^{-3}. Through algorithm development and validation analyses, these data form the principal foundation for ocean color satellite climate data record development. Here, we describe the cohesive OBPG system for field data analysis and satellite data product validation, from evaluation and distribution of data and results via SeaBASS to the integration of the data into NOMAD. Further, the latest updates to NOMAD are summarized and recent OBPG algorithm development results are presented.

2. SEABASS

SeaBASS includes data from over 1,500 field campaigns, contributed by 65 funded and volunteer researchers at 47 institutions in 14 countries (Werdell and Bailey 2002). Nearly 90% were collected by participants in the NASA SIMBIOS Program (Fargion et al. 2003). Data include measurements of marine optical properties, phytoplankton pigment concentrations, and other related oceanic and atmospheric observations. To facilitate consistency amongst data contributors and instruments, the OBPG specified *a priori* a series of *in situ* data requirements and sampling strategies that ensure the observations are acceptable for algorithm development and satellite validation (Mueller et al. 2003).



DATA FILE FORMAT

HEADER

```
/begin_header
/keyword(1)=value(1)
...
/date=YYYYMMDD
/time=HH:MM:SS[GMT]
/latitude=DD.DD[DEG]
/wind_speed=W[m/s]
...
/keyword(N)=value(N)
/end_header
```

DATA

```
value(1,1),...,value(1,N)
value(2,1),...,value(2,N)
value(3,1),...,value(3,N)
...
...
...
value(M,1),...,value(M,N)
```

QUALITY CONTROL

DIRECTORY TREE

```
/SeaBASS
└/NASA_GSFC
  ├/AMT
  │ ├/AMT1
  │ │ ├/archive
  │ │ └/documents
  │ └/AMT2
  │   ├/archive
  │   └/documents
  └/ROAVERRS
    ├/ROAVERRS_96
    │ ├/archive
    │ └/documents
    └/ROAVERRS_97
      ├/archive
      └/documents
```

DATABASE

CRUISE: CRUISE_ID, NAME, PI, ...

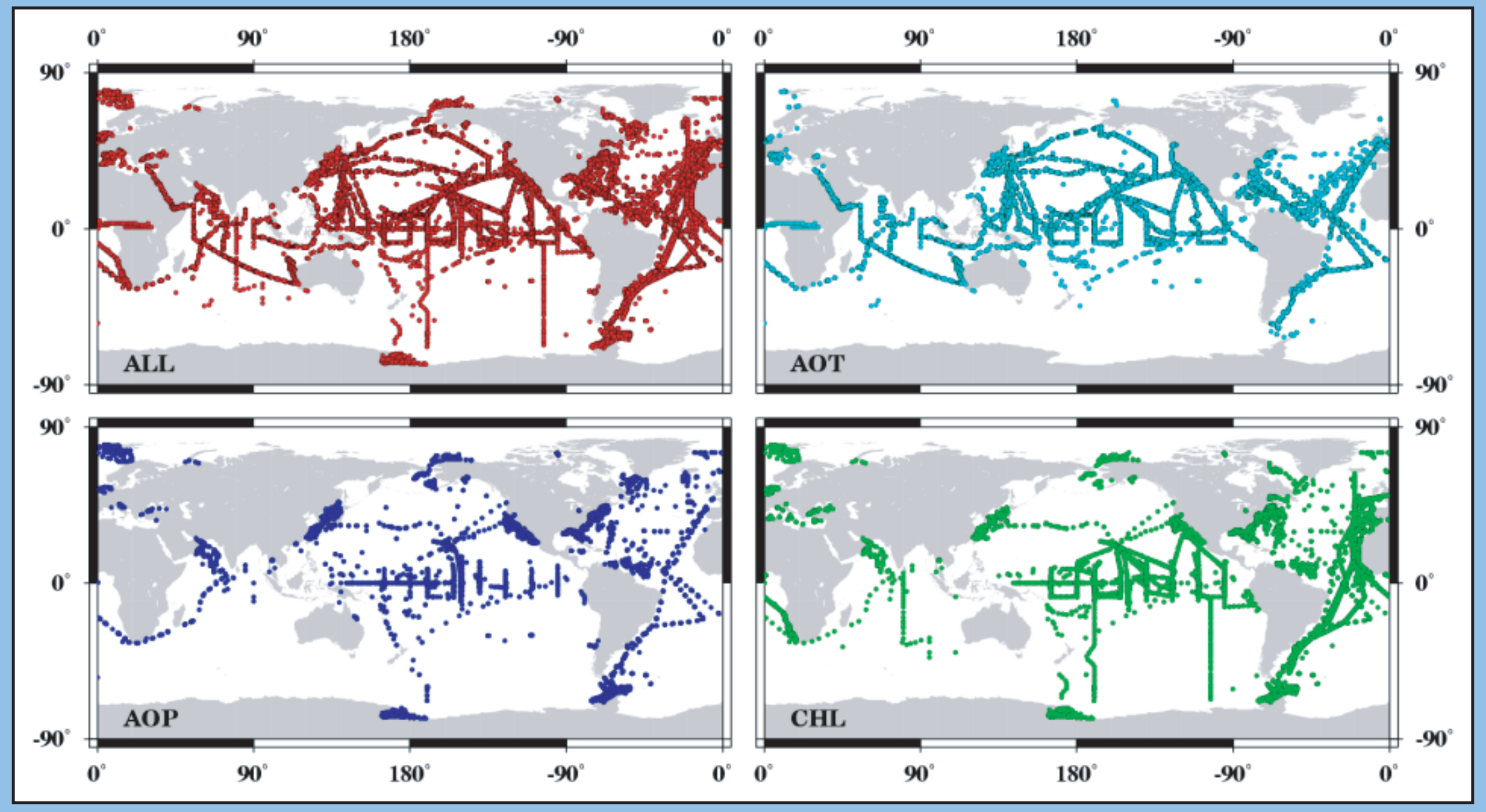
FILE: FILE_ID, PATH, CRUISE_ID, ...

METADATA: DATA_ID, DATE, LATITUDE, DEPTH, FILE_ID, ...

DATA: DATA_ID, NAME, VALUE, ...

WORLD WIDE WEB

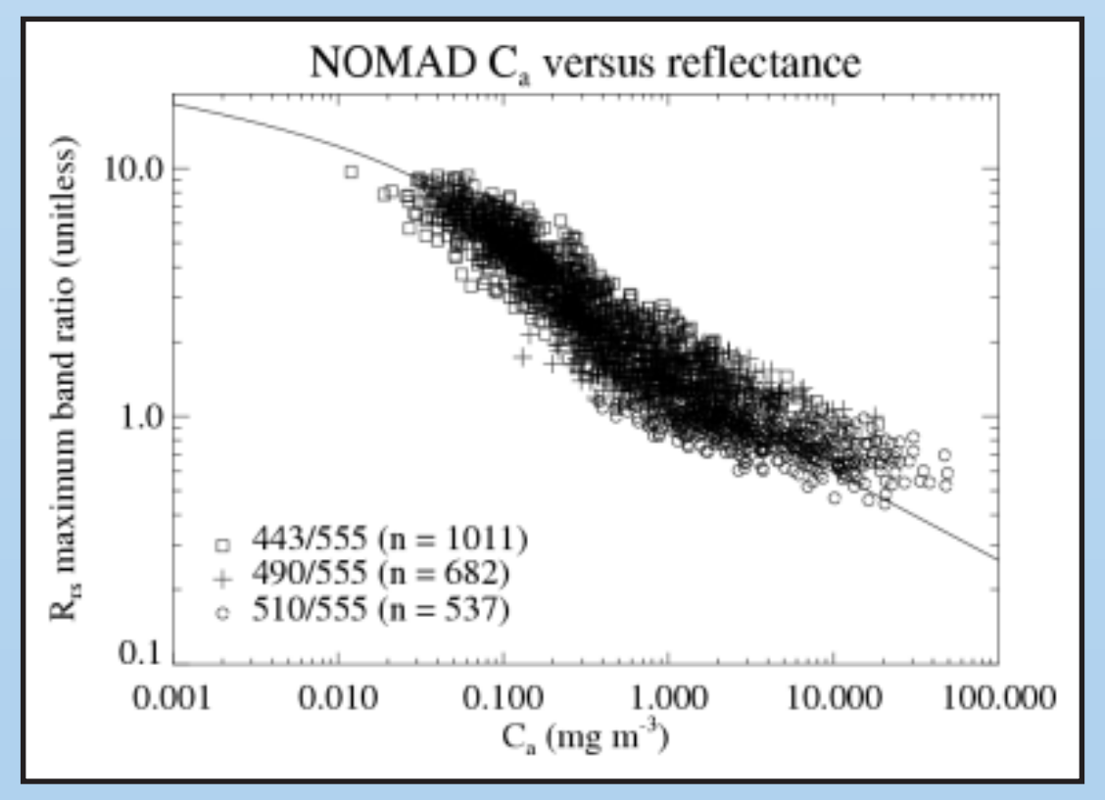
As indicated above, the architecture of SeaBASS consists of three tiers: geophysical data and metadata recorded in digital ASCII text files; a directory tree structure residing on a dedicated server at NASA/GSFC for storage of the data files; and, a relational database management system (RDBMS), built using the SQL Server product from Sybase, Inc., to catalog and distribute the data and files. Through the use of online search engines that interface with the RDBMS, the full bio-optical data set is queriable and available to authorized users via the World Wide Web at http://seabass.gsfc.nasa.gov.



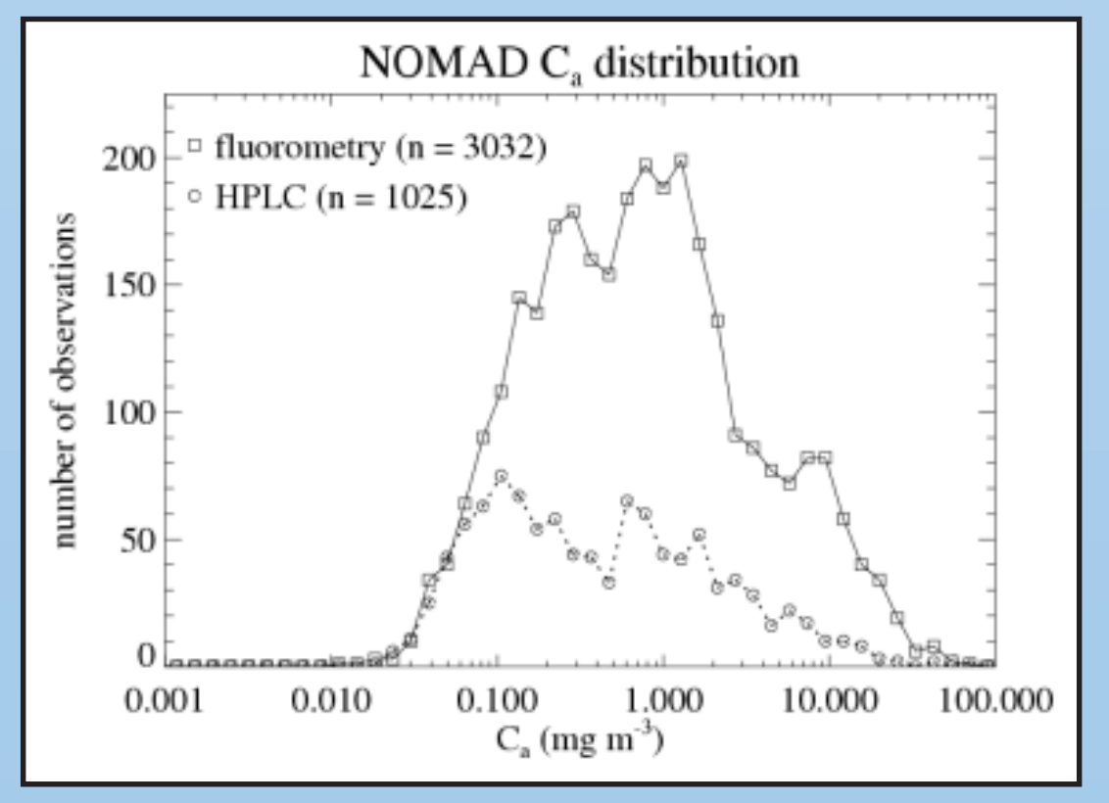
SeaBASS includes a nearly global distribution of data products, with the exception of a notable paucity of observations in the Southern Ocean. Clockwise from upper left: all data stations (**ALL**); aerosol optical thicknesses only (**AOT**); C_a only (**CHL**); and, radiometric observations only (**AOP**).

3. NOMAD

To expedite community bio-optical algorithm development, the NASA OBPG generated a standalone *in situ* data set of remote-sensing relevant products, named NOMAD, using the data archived in SeaBASS (Werdell and Bailey 2005). It consists of consistently processed coincident surface observations of radiometric spectra, bulk water temperatures and salinities, C_a, and associated metadata, such as the time and location of data collection. NOMAD is publicly available via http://seabass.gsfc.nasa.gov/cgi-bin/nomad.cgi.



NOMAD C_a versus reflectance



NOMAD C_a distribution

λ-nm	%	HERITAGE SENSOR
405	2.5	MOS
411	99.5	OCTS, SeaWiFS, MODIS, MERIS, VIIRS
443	99.9	CZCS, OCTS, MOS, SeaWiFS, MODIS, MERIS, VIIRS
455	15.8	
465	3.8	
489	100.0	OCTS, MOS, SeaWiFS, MODIS, MERIS, VIIRS
510	75.8	SeaWiFS, MERIS
520	42.6	CZCS, OCTS, MOS
530	34.9	MODIS
550	21.7	CZCS, MODIS
555	70.0	SeaWiFS, VIIRS
560	21.7	MERIS
565	46.2	OCTS
570	18.9	MOS
590	13.5	
619	17.2	MOS, MERIS
625	43.5	
665	59.0	MODIS, MERIS
670	27.4	CZCS, OCTS, SeaWiFS, VIIRS
683	45.5	MOS, MERIS

Above, we graphically describe the current geophysical distribution of data included in NOMAD. Upper left: remote-sensing reflectance, R_{rs}, maximum band ratios as a function of C_a. For reference, the operational SeaWiFS C_a algorithm (OC4 version 4; O'Reilly et al. 2000), which exploits this relationship, is shown as the solid line. Lower left: the frequency distribution of fluorometrically- and HPLC-derived C_a. Center: the spectral resolution of the radiometric data, including the frequency of occurance of each wavelength within the data set and the heritage sensors possessing (approximately) each wavelength.

- Generic form of a sea surface reflectance model:

$$R_{rs}(\lambda) \sim \frac{b_b(\lambda)}{a(\lambda)}, \text{ where}$$

$$a(\lambda) = a_w(\lambda) + a_d(\lambda) + a_g(\lambda) + a_\phi(\lambda), \text{ and}$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda), \text{ and}$$

w, d, g, ϕ, and p indicate water, non-algal detritus, dissolved organics, phytoplankton, and total particles, respectively.

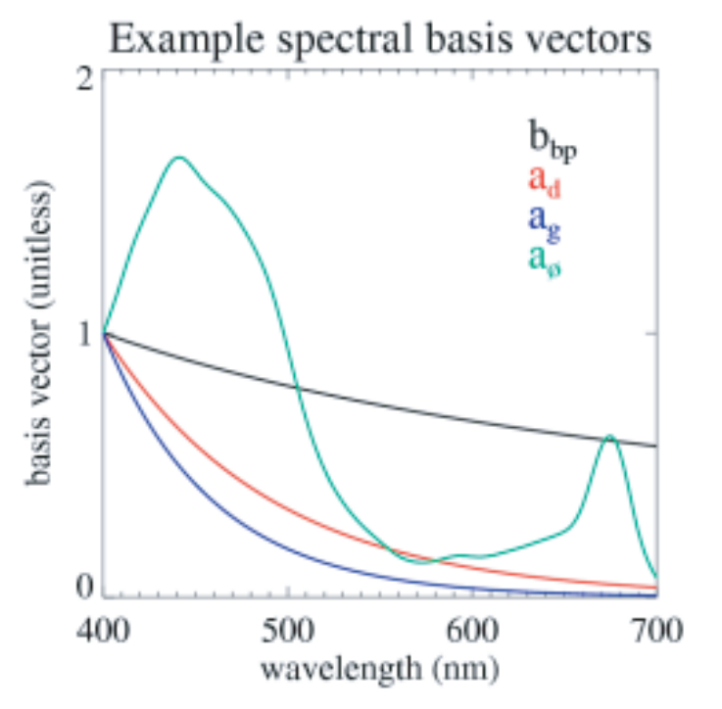
- The water components are known.
- NOMAD now includes the red parameters.

- In the inversion, the a and b_b terms are described as a function of their *spectral (basis) shapes*, which are considered known, and their *magnitudes*, which are the solutions:

$$x(\lambda) = M_x\,\bar{\mathbf{x}}(\lambda), \text{ where } x \text{ is any product, or,}$$

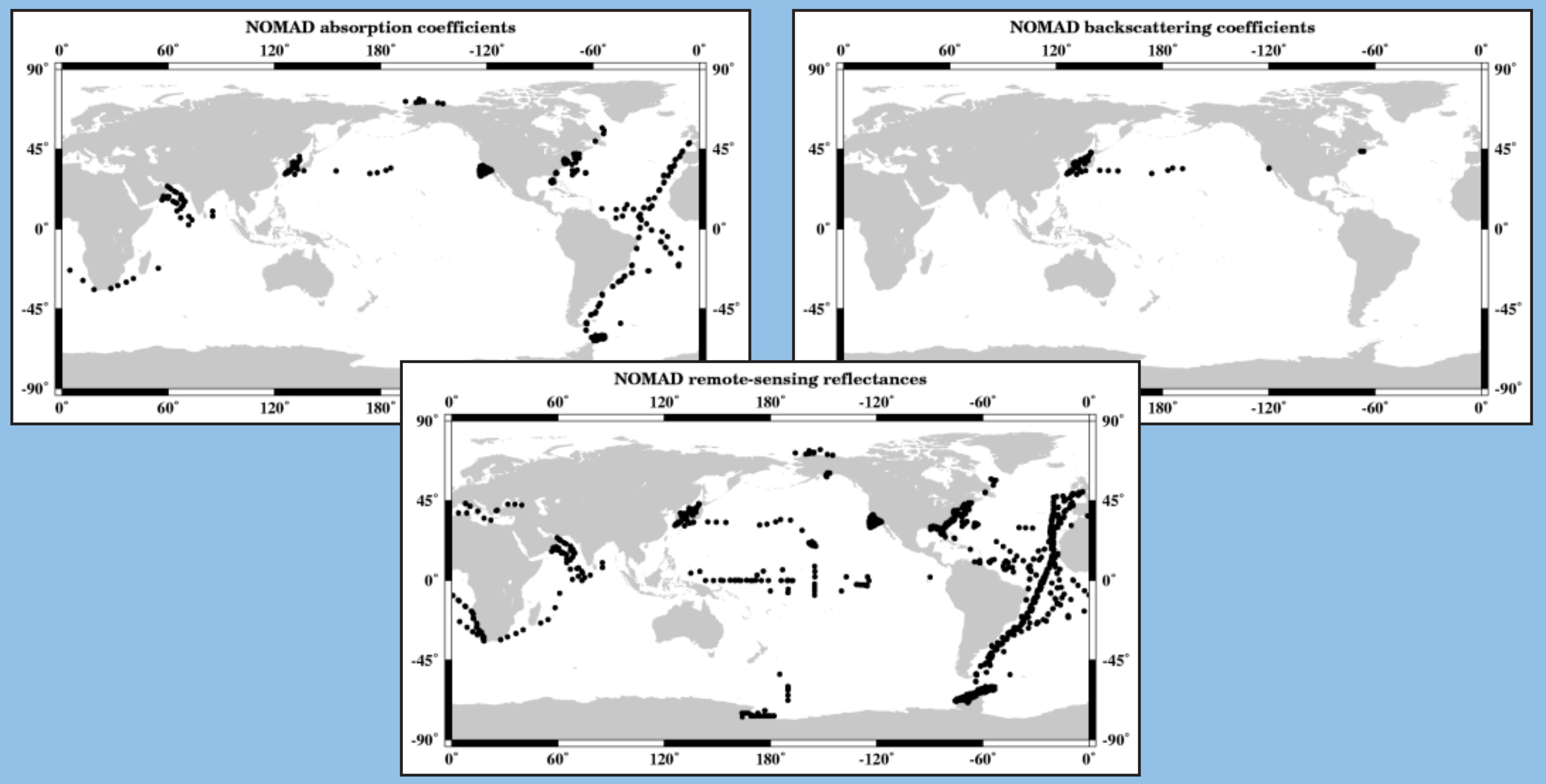
$$b_{bp}(\lambda) = M_{bp}\,\lambda^{-1}, \text{ as a specific example.}$$

- While the inclusion of these spectral parameters in NOMAD supports algorithm validation analyses, it also expedites regional and global tuning of such algorithms and their input basis vectors.



Example spectral basis vectors

Initially, NOMAD radiometric data consisted solely of water-leaving radiances, $L_w(\lambda)$, surface irradiances, $E_s(\lambda)$, and downwelling vertical attenuation coefficients, $K_d(\lambda)$. Note, the ratio of the first two gives R_{rs} (= L_w/E_s). Recently, NOMAD was expanded to include spectral absorption, $a(\lambda)$, and backscattering, $b_b(\lambda)$, coefficients to support the development and refinement of sea surface reflectance inversion models, as highlighted in the text box above. To date, $a(\lambda)$ in NOMAD are derived exclusively from spectroscopy.



NOMAD absorption coefficients

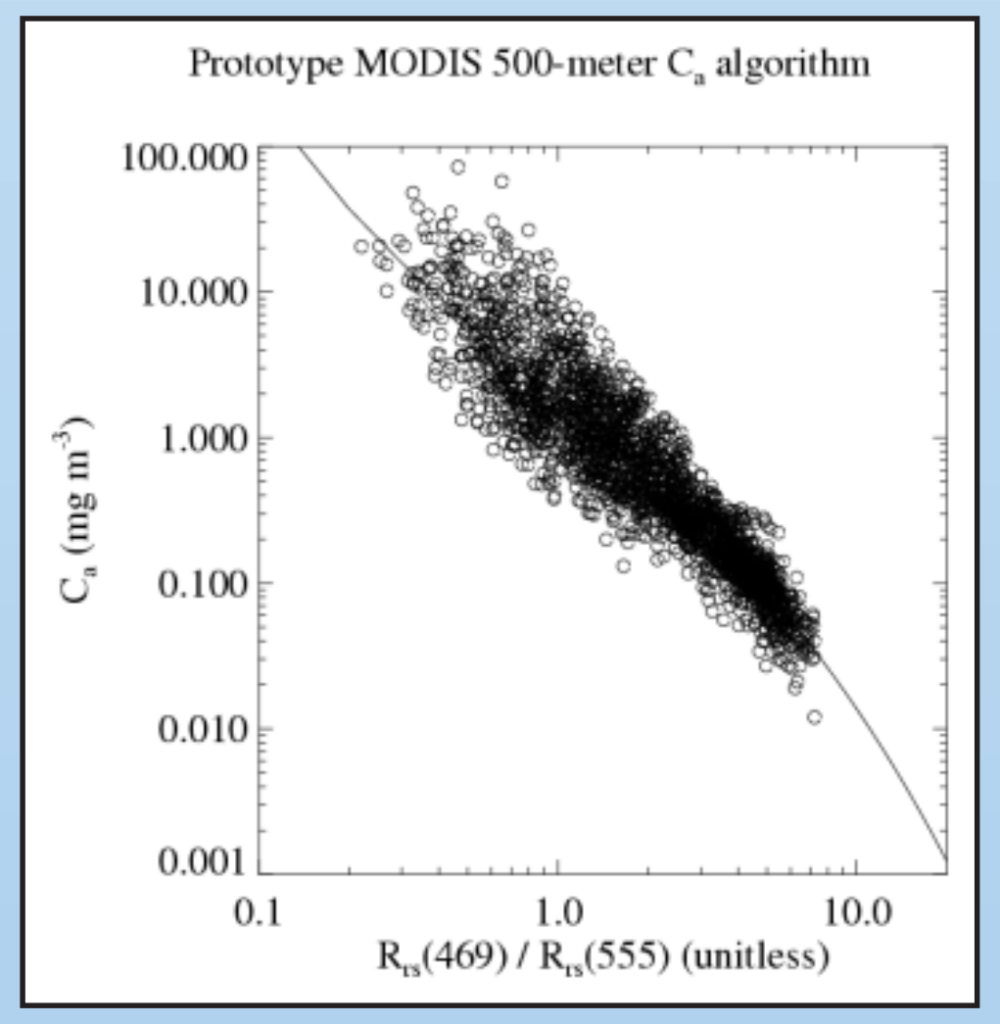
NOMAD backscattering coefficients

NOMAD remote-sensing reflectances

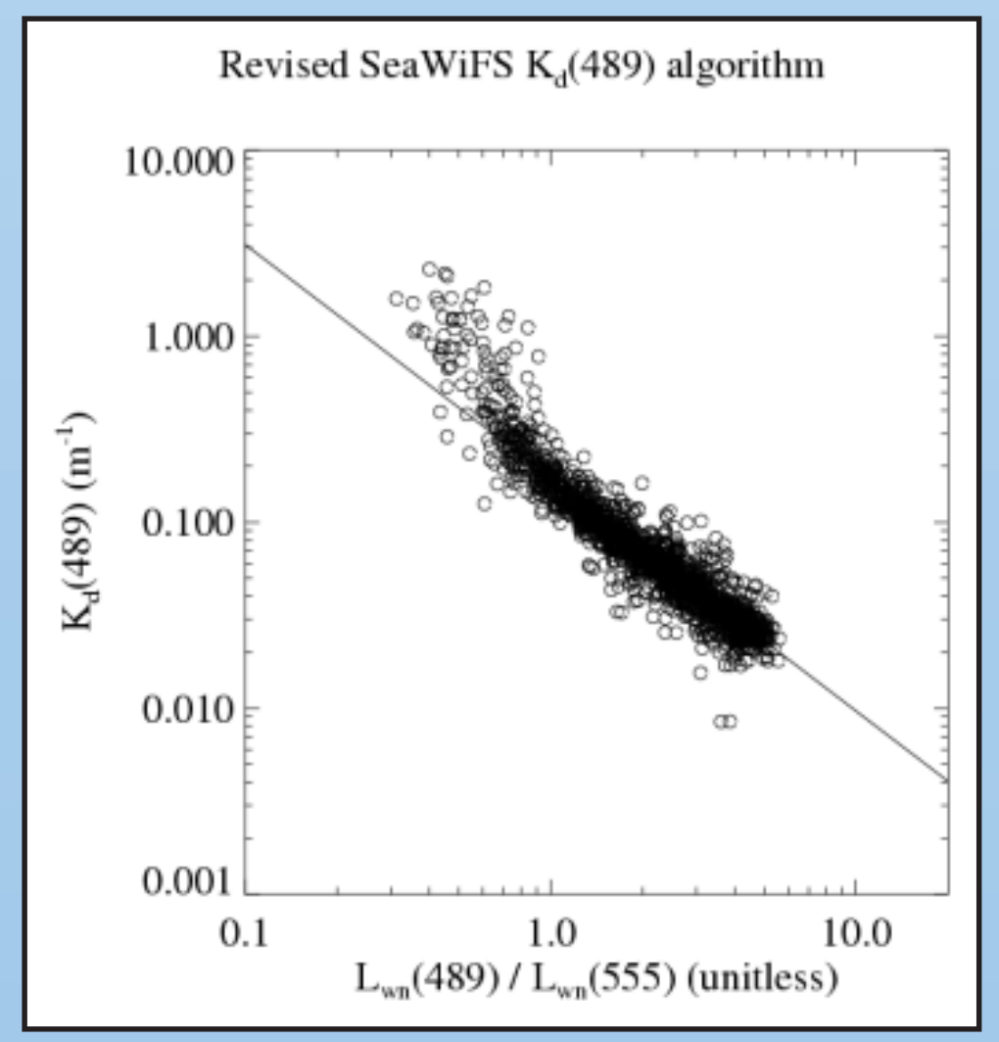
While NOMAD $R_{rs}(\lambda)$ are reasonably globally represented, coincident measurements of $a(\lambda)$ and $b_b(\lambda)$ are currently sparse. Note that a C_a value accompanies each record.

4. ALGORITHMS & VALIDATION

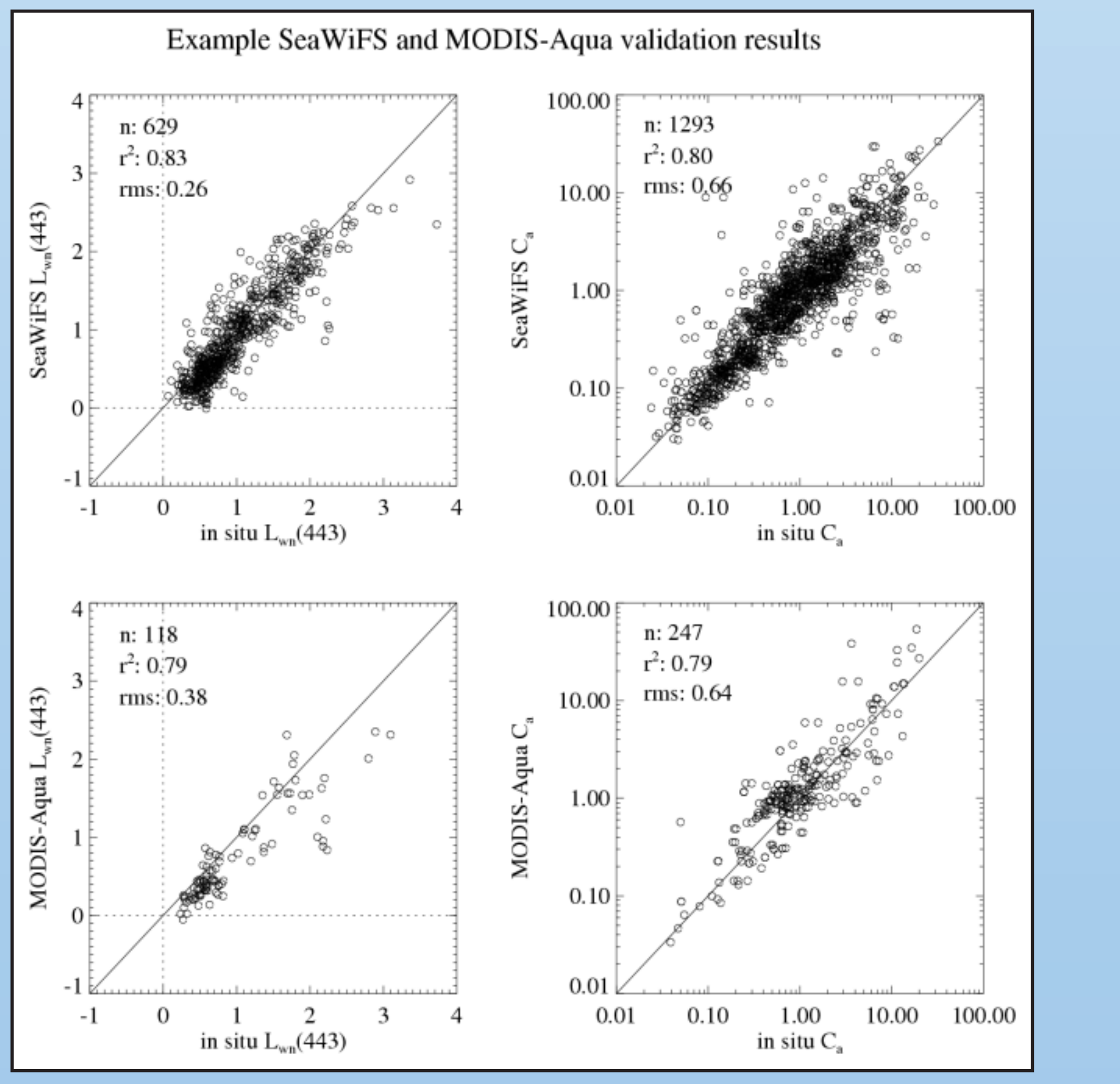
The NASA OBPG employs NOMAD in both its bio-optical algorithm development and satellite data product validation activities, as shown below. The OBPG also evaluates community-supplied algorithms using NOMAD, when appropriate.



Prototype MODIS 500-meter C_a algorithm



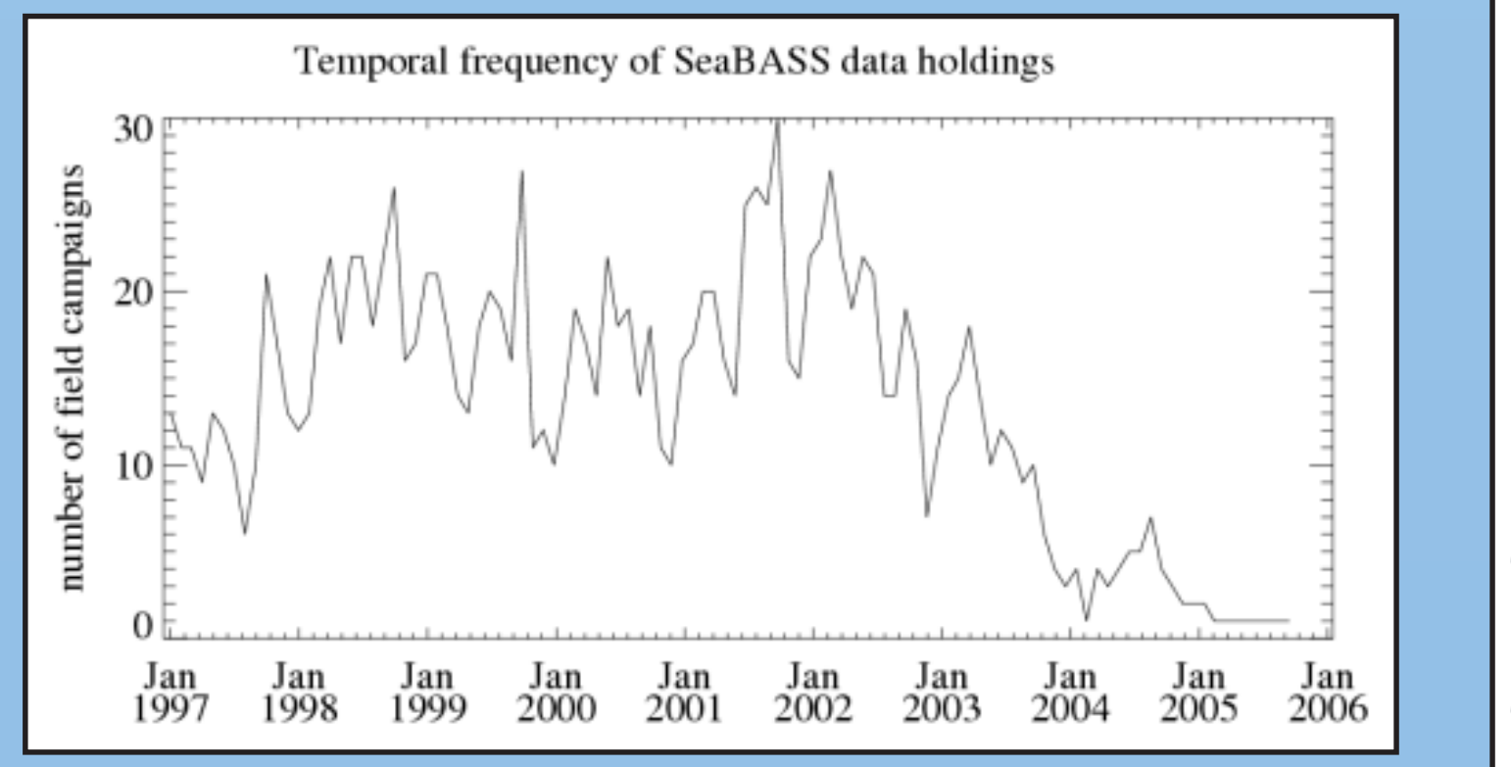
Revised SeaWiFS $K_d(489)$ algorithm



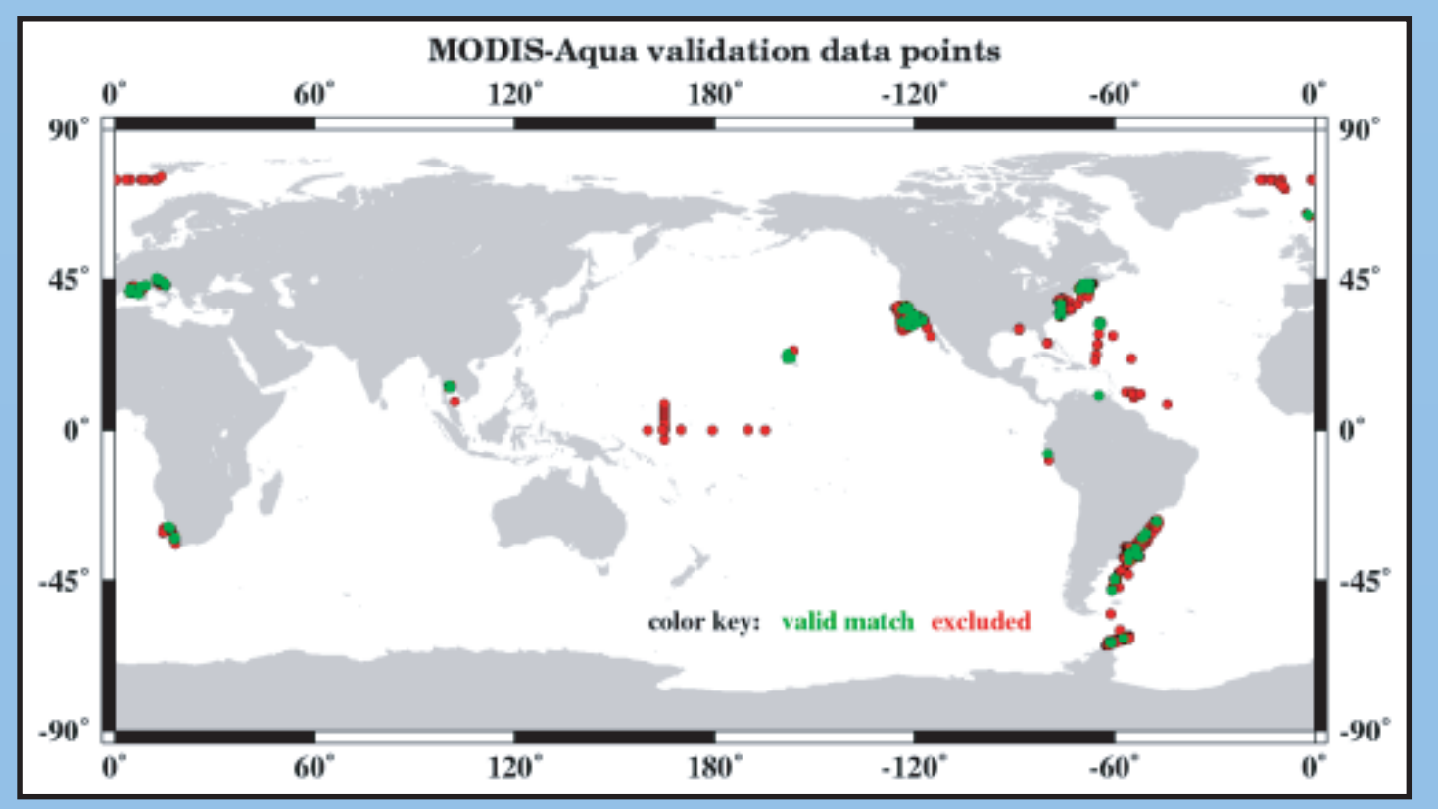
Example SeaWiFS and MODIS-Aqua validation results

Results from two recent algorithm development activities are shown on the left. Top: an R_{rs} function, analogous to OC4, designed to estimate C_a using the MODIS 500-meter bands at 469 and 555-nm. As $R_{rs}(469)$ is uncommon in NOMAD, it was estimated via interpolation of four surrounding bands. Bottom: the updated form of the operational SeaWiFS K_d algorithm (O'Reilly et al. 2000). The OBPG used NOMAD to generate both empirical relationships.

Validation results for $L_{wn}(443)$ and C_a are presented on the right for SeaWiFS (reprocessing 5.1, July 2005) and MODIS-Aqua (reprocessing 1.1, August 2005). The OBPG operationally evaluates the performance of both sensors using NOMAD as ground truth. Satellite validation results are routinely posted online at http://seabass.gsfc.nasa.gov for community comment and perusal.



Temporal frequency of SeaBASS data holdings



MODIS-Aqua validation data points

The volume of data contributions to SeaBASS has steadily declined since the end of the SIMBIOS era in 2003, as indicated graphically above. This absence of *in situ* data hinders the data product validation of the most recently launched satellite instruments, such as MODIS-Aqua, whose ocean color data stream began in June 2004. The lack of geographic variability of the observations, for example, limits the validation of MODIS-Aqua to specific regions and a narrow range of biogeophysical conditions.

5. ACKNOWLEDGEMENTS & REFERENCES

This work was funded by the NASA Earth Observing System (EOS)/MODIS and NASA Ocean Biogeochemistry Programs. The authors are grateful to G. Feldman, G. Fargion, and all OBPG Staff members for their regular support and assistance in this effort.

Fargion, G.S. and 9 co-authors, 2003: *Proceedings SPIE 5155*, 49-60.
Mueller, J.L. and 17 co-authors, 2003: *NASA Tech. Memo. 2003-211621/Rev4-Vol3*, 78 pp.
O'Reilly, J.E. and 24 co-authors, 2000: *NASA Tech. Memo. 2000-206892 Vol. 11*, 49 pp.
Werdell, P.J. and Bailey, S.W. , 2002: *NASA Tech. Memo. 2002-211617*, 45 pp.
Werdell, P.J. and Bailey, S.W. , 2005: *Remote Sensing of Environment*, 98, 122-140.

All references are available online at http://oceancolor.gsfc.nasa.gov.