Ocean Colour Algorithms

Mark Dowell
Joint Research Centre
mark.dowell@jrc.ec.europa.eu

Session Outline

4 Sections

1. Historical foundations and development of Ocean Colour Algorithms
2. Basis for implementing semi-analytical algorithms
3. Alternative inversion methods and comparisons
4. Algorithm implementation issues, regional and class based algorithms

Disclaimers!

- I do not have an algorithm of my own that I will be pushing (but I do have opinions)
- The lectures can not be exhaustive (i.e. cover all algorithm types) – focus on common truths/lessons
- There may be some material which is already known by some – but hopefully there will be something useful for all

Overarching issues (e.g.)

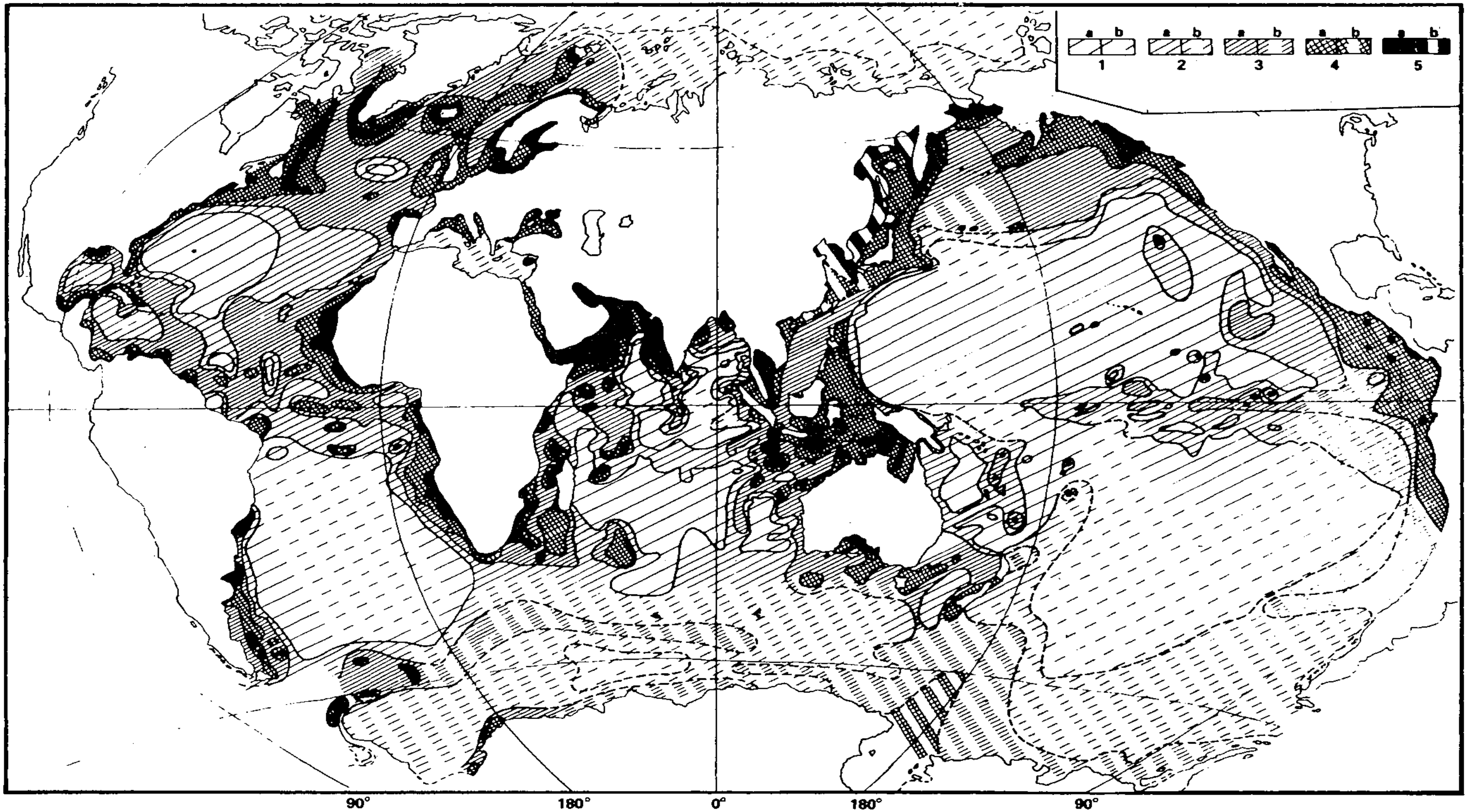
- The evolution of ocean colour algorithms has resulted from improvements on our knowledge on bio-optics but importantly also on the availability of relevant data
- Beware of claims of a purely analytical algorithm, there is some degree of empiricism in all algorithms & we shouldn't be ashamed of this
- Use of the data is important – i.e. clear separation between "parameterisation" datasets and validation datasets
- Understanding (ideally performing) in-situ measurements will help you to understand the caveats and applicability of algorithms you may develop
- Beware of "assumptions" made by algorithms e.g. nLw products that have been constrained to an in water model during atmospheric correction
- The scientific state-of-the-art is not always compatible with routine

Broad topics not covered in detail

- Data Merging
- Uncertainties – Doerffer lectures
- IOP variability – Lee lectures

Section 1

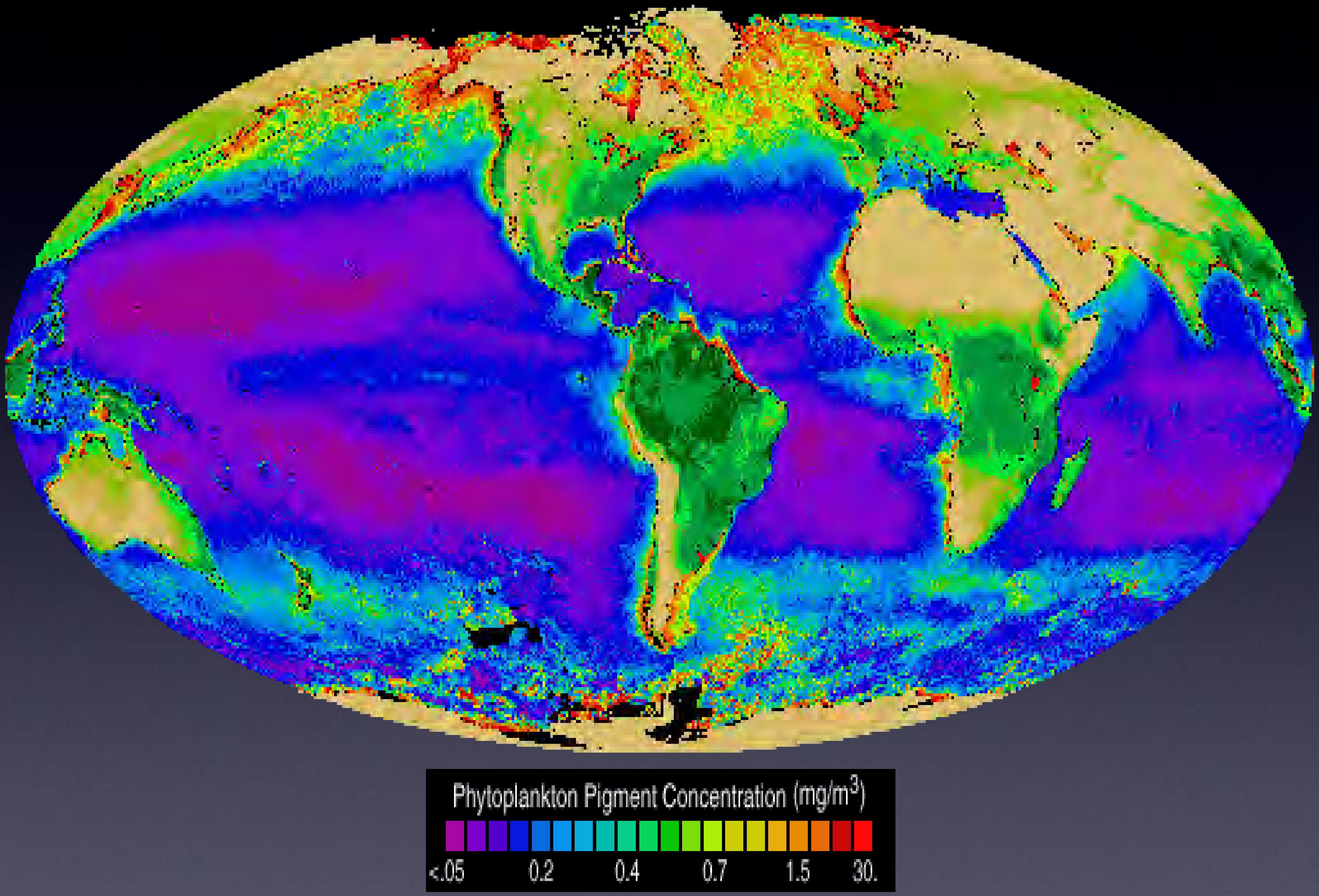
Historical foundations and development of Ocean Colour Algorithms



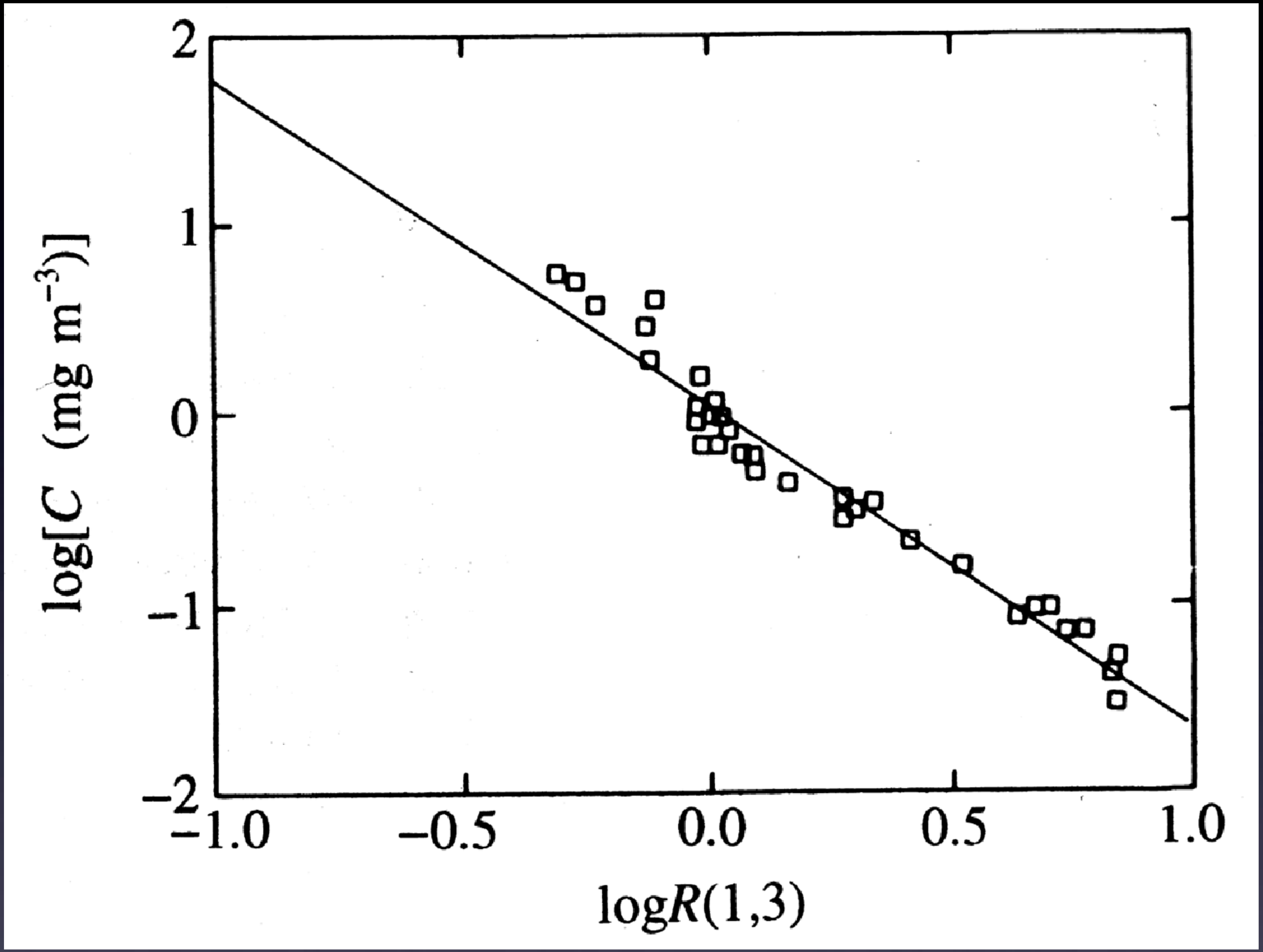
URE 1 Distribution of primary production in the World Ocean. Units are in mg of C per m^2 per day. (1) Less than 100; (2) 100–150; 150–250; (4) 250–500; (5) more than 500. *a* = data from direct ^{14}C measurements; *b* = data from phytoplankton biomass, hydrogen, oxygen saturation.

Olga Koblentz-Mishke 1970

Coastal Zone Color Scanner (CZCS) Global Climatology (Nov. 1978-Jun. 1986)



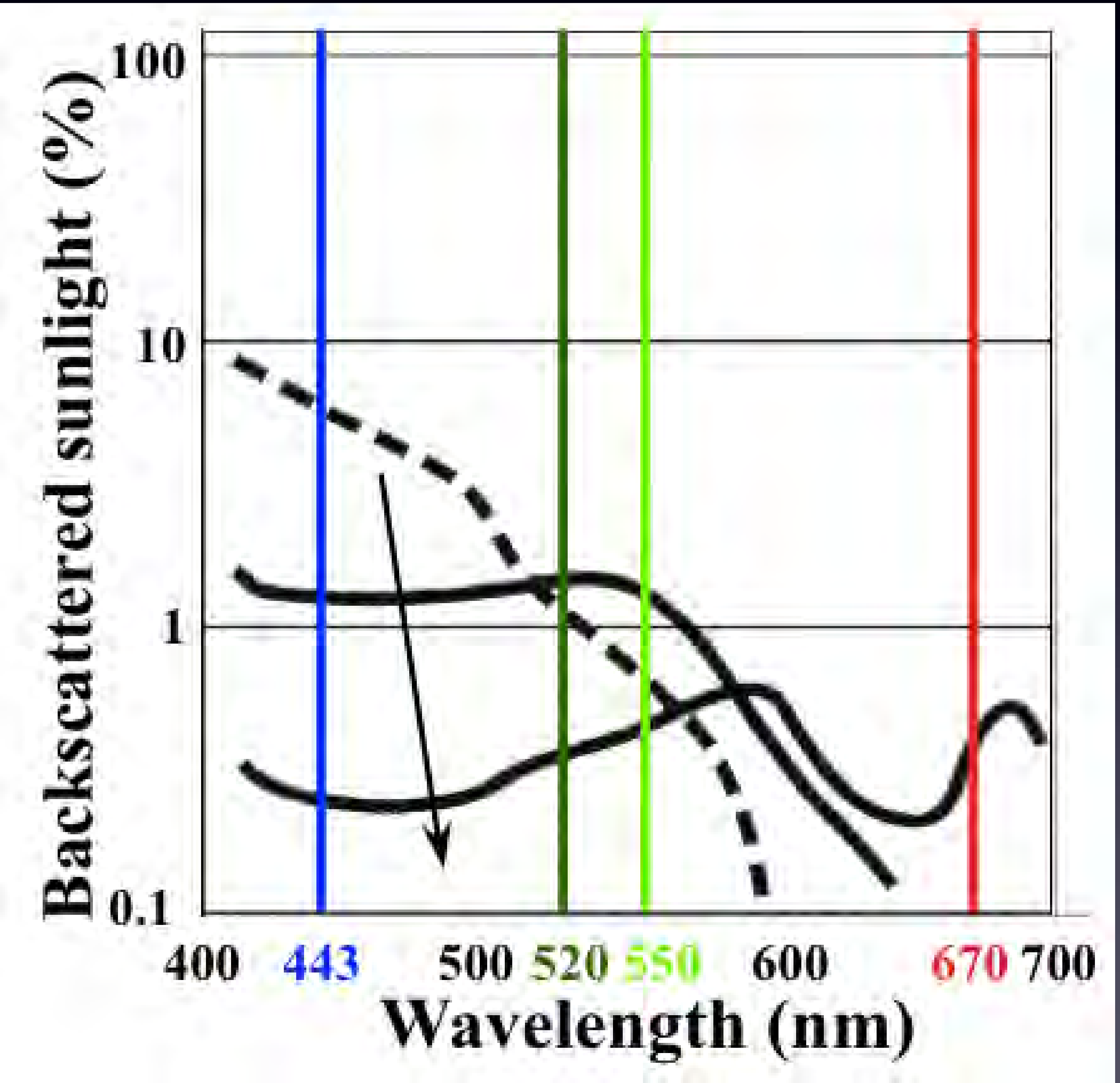
$R(1,3) = L_w(B_1{=}443)/L_w(B_3{=}550)$ vs Chl

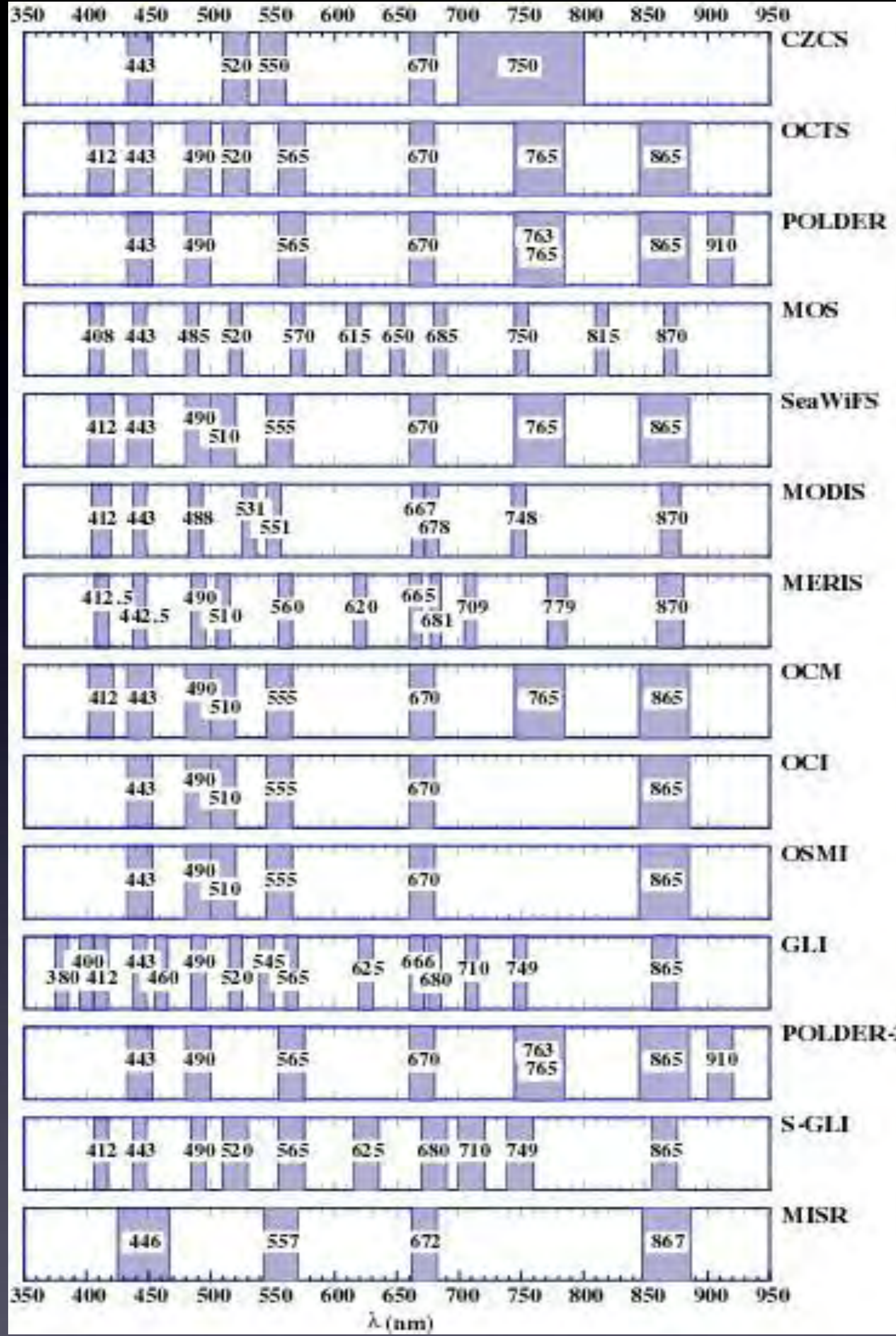


Note: only 33 data points were initially available!

Suggests the band-ratio model: $\log_{10}(\mathrm{Chl}) = C_1 + C_2 \log_{10}[L_w(443)/L_w(550)]$
C_1 and C_2 are the model parameters whose values are determined by the data

Basic Ocean Colour “paradigm”



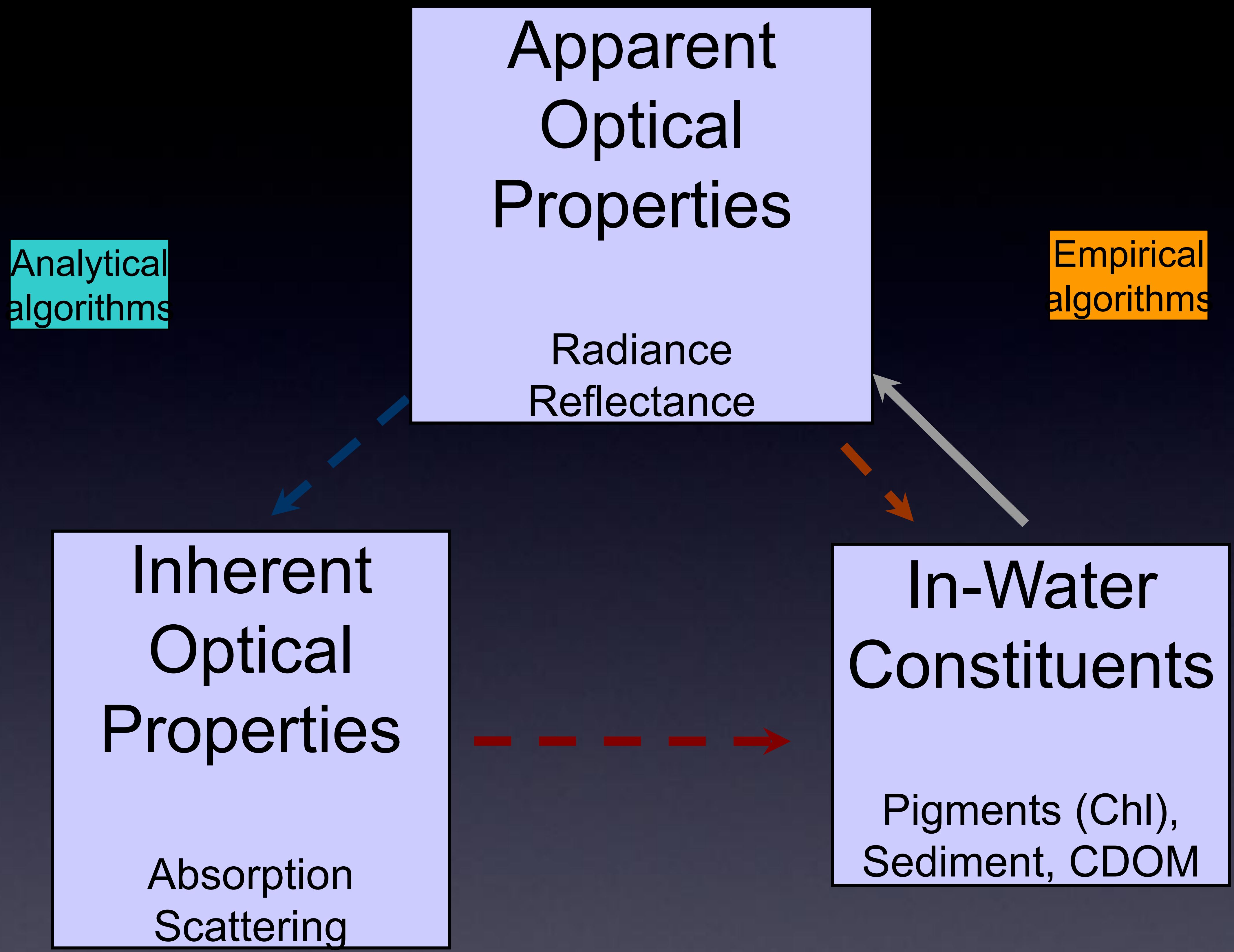


CZCS
443 520 550 670 750
OCTS
412 443 490 520 565 670 765 865
POLDER
443 490 565 670 763 765 865 910
MOS
408 443 485 520 570 615 650 685 750 815 870
SeaWiFS
412 443 490 510 555 670 765 865
MODIS
412 443 488 531 551 667 678 748 870
MERIS
412.5 442.5 490 510 560 620 665 681 709 779 870
OCM
412 443 490 510 555 670 765 865
OCI
443 490 510 555 670 865
OSMI
443 490 510 555 670 865
GLI
380 400 412 443 460 490 520 545 565 625 666 680 710 749 865
POLDER-2
443 490 565 670 763 765 865 910
S-GLI
412 443 490 520 565 625 680 710 749 865
MISR
446 557 672 867
350 400 450 500 550 600 650 700 750 800 850 900 950
λ (nm)

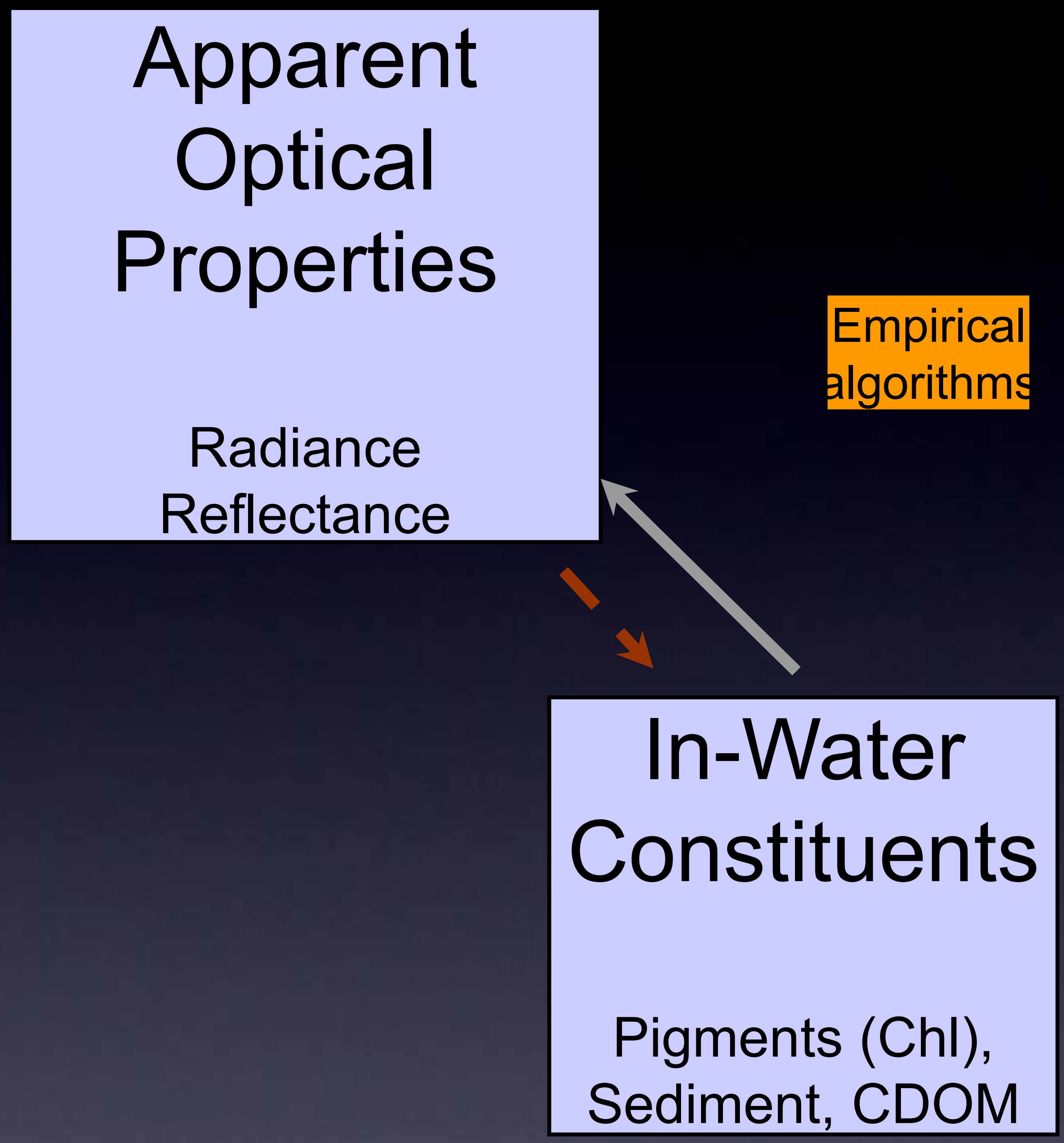
Ocean colour algorithms

- Two kinds:

1. Empirical: often (but not always) chlorophyll only
2. Semi-analytic: chlorophyll, CDOM, particulates (but also more PSD, PFTs)



Local parameterization for coastal & inland waters


Apparent Optical Properties
Radiance
Reflectance
Empirical algorithms
In-Water Constituents
Pigments (Chl),
Sediment, CDOM

Examples of Band-Ratio Algorithms

SeaWiFS OC4 for Chl:

$X = \log_{10}\{\max[R_{rs}(443)/R_{rs}(555), R_{rs}(490)/R_{rs}(555), R_{rs}(510)/R_{rs}(555)]\}$

$Chl = 10\hat{}(0.366 - 3.067X + 1.930X^2 + 0.649X^3 - 1.532X^4)$

MODIS for $K_d(490)$:

$X = L_w(488)/L_w(551)$

$K_d(490) = 0.016 + 0.156445X\hat{}(-1.5401)$

MODIS for $a_{CDOM}(400)$ and $a_{phy}(675)$:

$r_{15} = \log_{10}[R_{rs}(412)/R_{rs}(551)]$

$r_{25} = \log_{10}[R_{rs}(443)/R_{rs}(551)]$

$r_{35} = \log_{10}[R_{rs}(488)/R_{rs}(551)]$

$a_{CDOM}(400) = 1.5*10\hat{}(-1.147 + 1.963r_{15} - 1.01r_{15}^2 - 0.856r_{25} + 1.02r_{25}^2)$

$a_{phy}(675) = 0.328\ [\ 10\hat{}(-0.919 + 1.037r_{25} - 0.407r_{25}^2 -$
$3.531r_{35} + 1.702r_{35}^2 - 0.008)]$

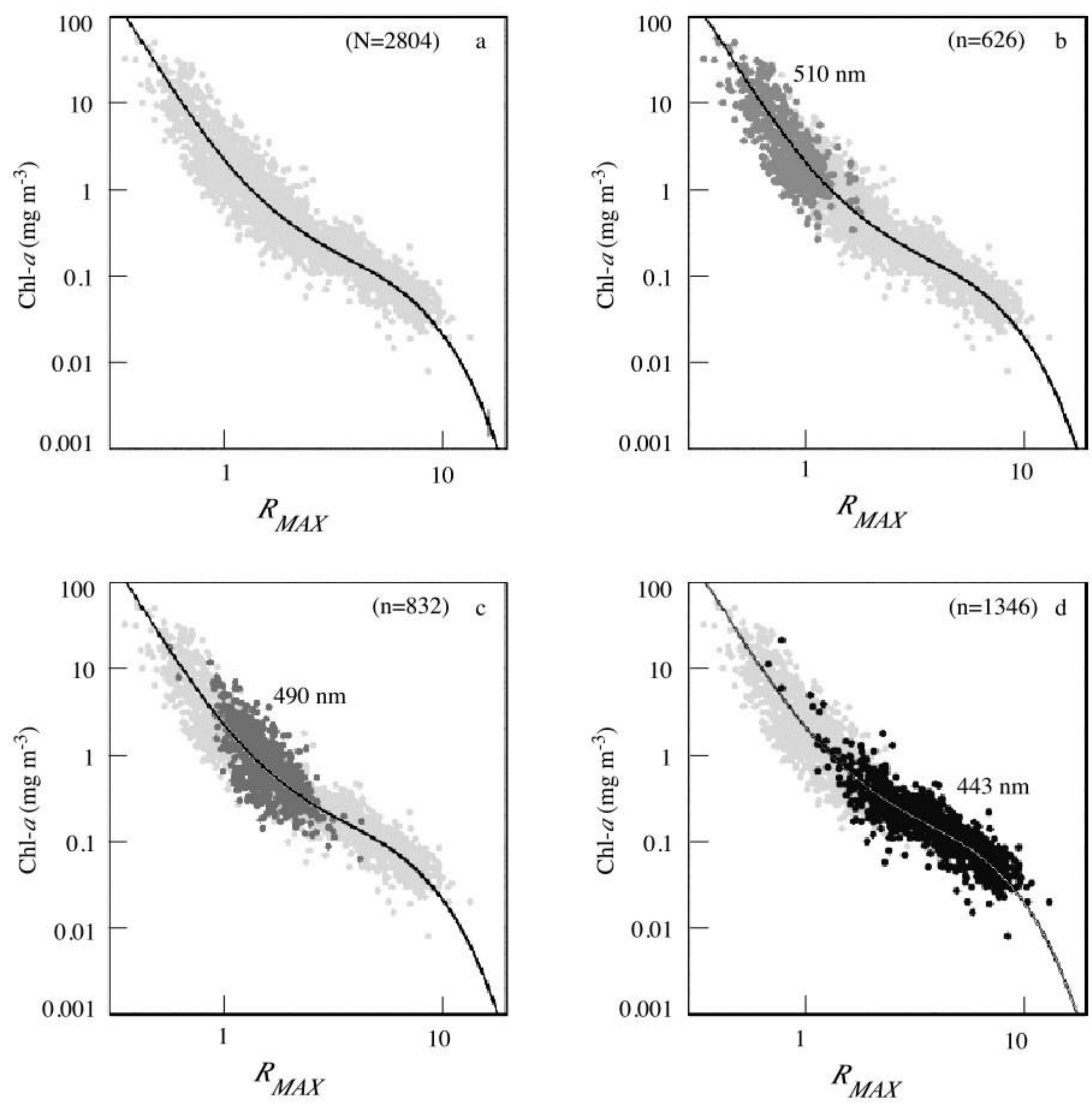
and so on, for dozens more….

SeaWiFS empirical OC4 algorithm for Chl-a; Called a *maximum-band ratio alg.*

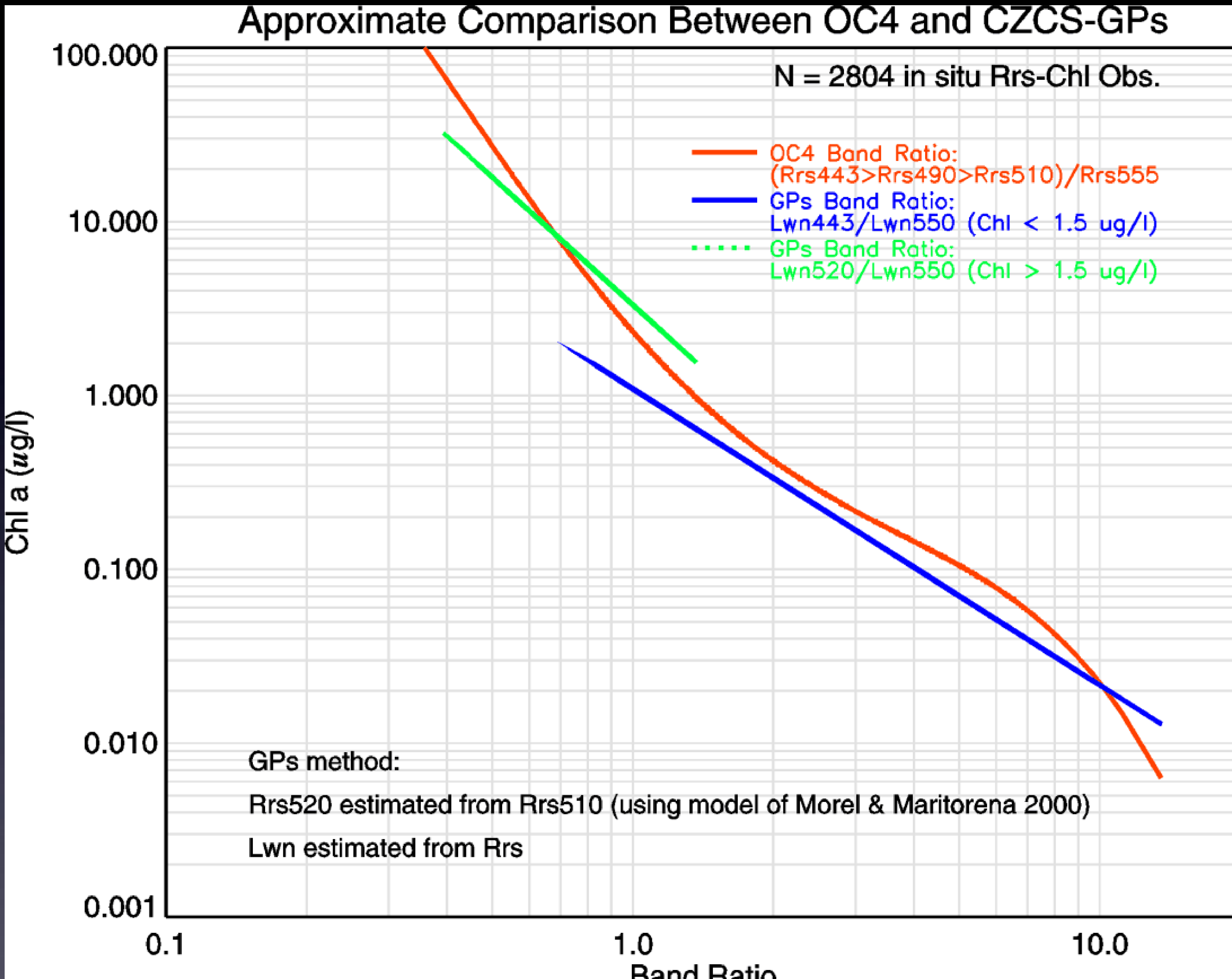
$$R_{MAX} = \text{Maximum of } [R_{rs}\text{-ratio}(443/555,\ 490/555,\ 510/555)]$$

$$R_L = \log_{10}(R_{MAX})$$

$$\log_{10}(C_a) = 0.366 - 3.067 R_L + 1.930 R_L^2 + 0.649 R_L^3 - 1.532 R_L^4$$

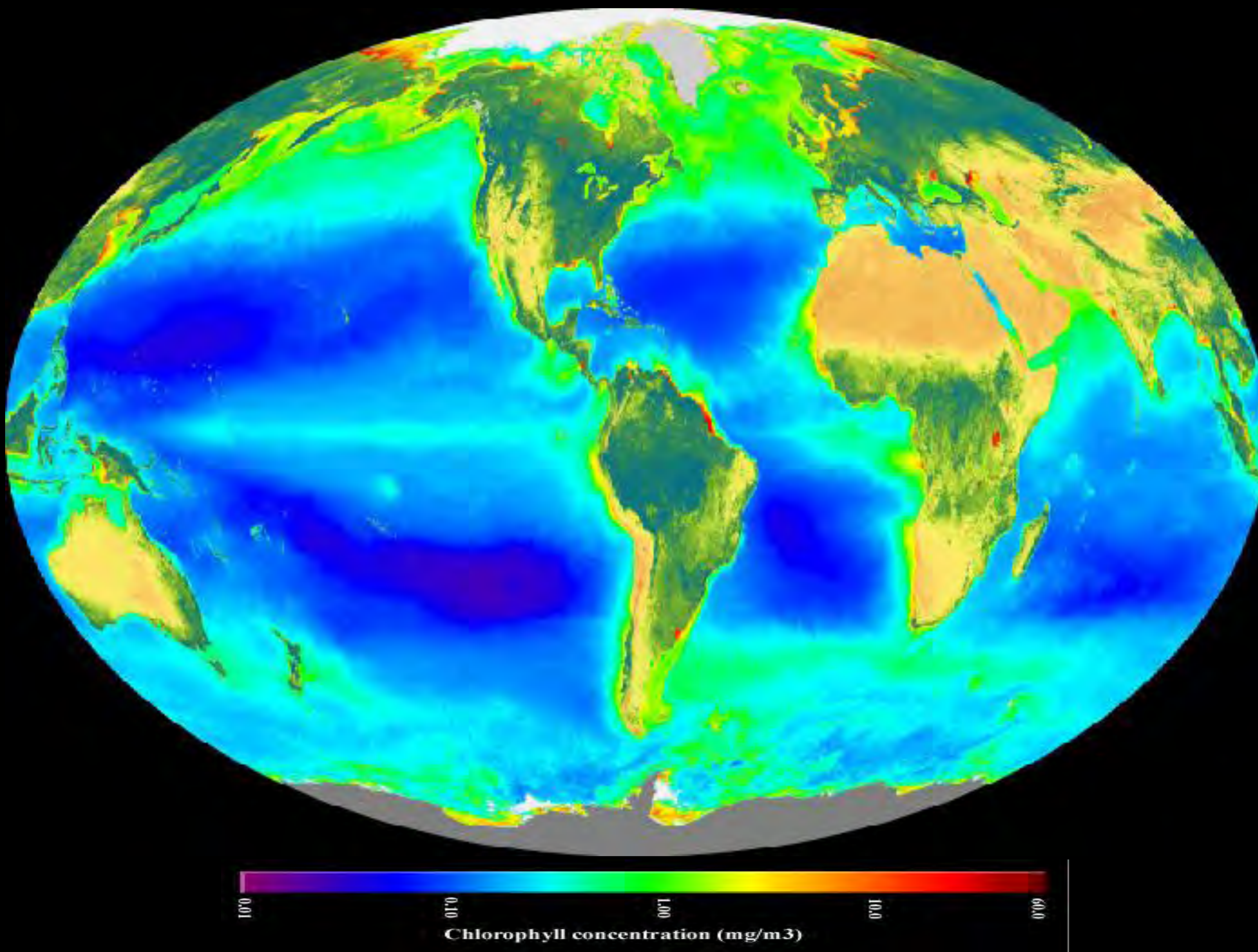


(N=2804) a
(n=626) b
510 nm
(n=832) c
490 nm
(n=1346) d
443 nm
Chl-a (mg m-3)
R MAX
100
10
1
0.1
0.01
0.001
1
10



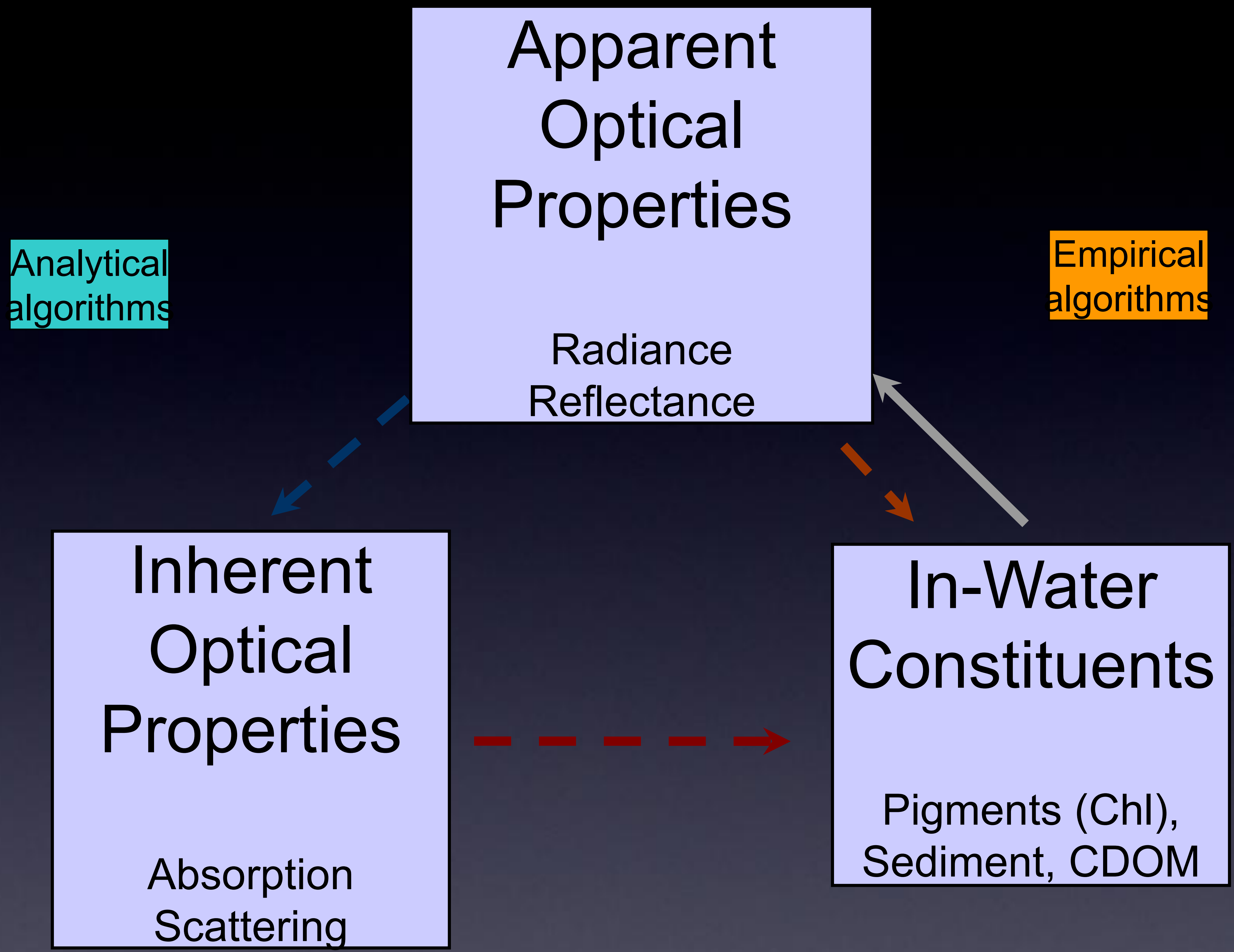
Approximate Comparison Between OC4 and CZCS-GPs
N = 2804 in situ Rrs-Chl Obs.
OC4 Band Ratio:
(Rrs443>Rrs490>Rrs510)/Rrs555
GPs Band Ratio:
Lwn443/Lwn550 (Chl < 1.5 ug/l)
GPs Band Ratio:
Lwn520/Lwn550 (Chl > 1.5 ug/l)
GPs method:
Rrs520 estimated from Rrs510 (using model of Morel & Maritorena 2000)
Lwn estimated from Rrs
Chl a (ug/l)
100.000
10.000
1.000
0.100
0.010
0.001
0.1
1.0
10.0
Band Ratio

SeaWiFS composite image (1997-2000)



Section 2

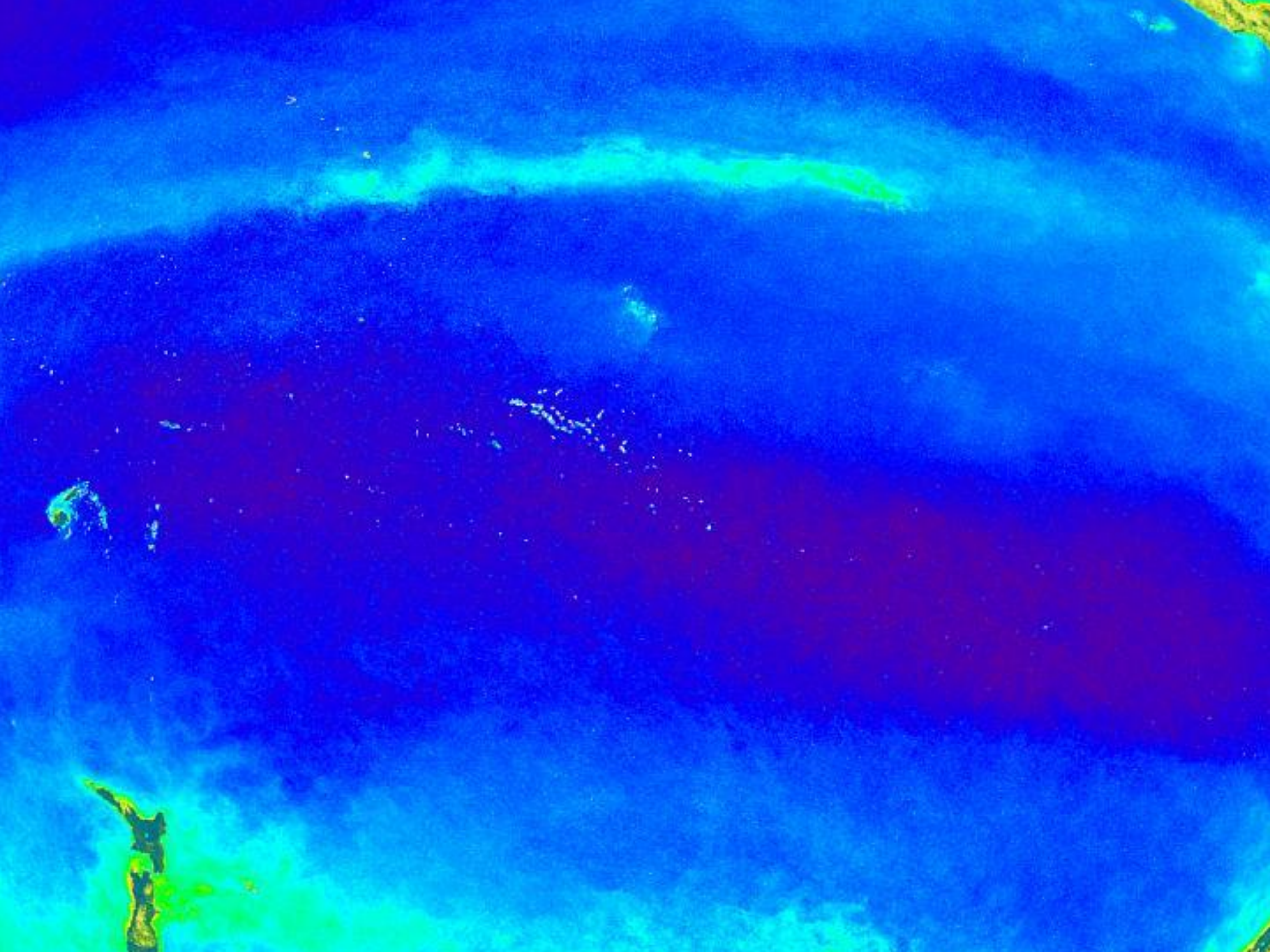
Basis for implementing semi-analytical algorithms



Apparent Optical Properties
Radiance
Reflectance
Analytical algorithms
Empirical algorithms
Inherent Optical Properties
Absorption
Scattering
In-Water Constituents
Pigments (Chl), Sediment, CDOM

For most ocean waters, phytoplankton are the only substance affecting the color of the water. These waters are called "Case 1" waters.

Case 1 waters are those waters where phytoplankton and covarying decay products are the only substances affecting the optical properties of the water. This is true almost everywhere except where there is some influence of land (or bottom).



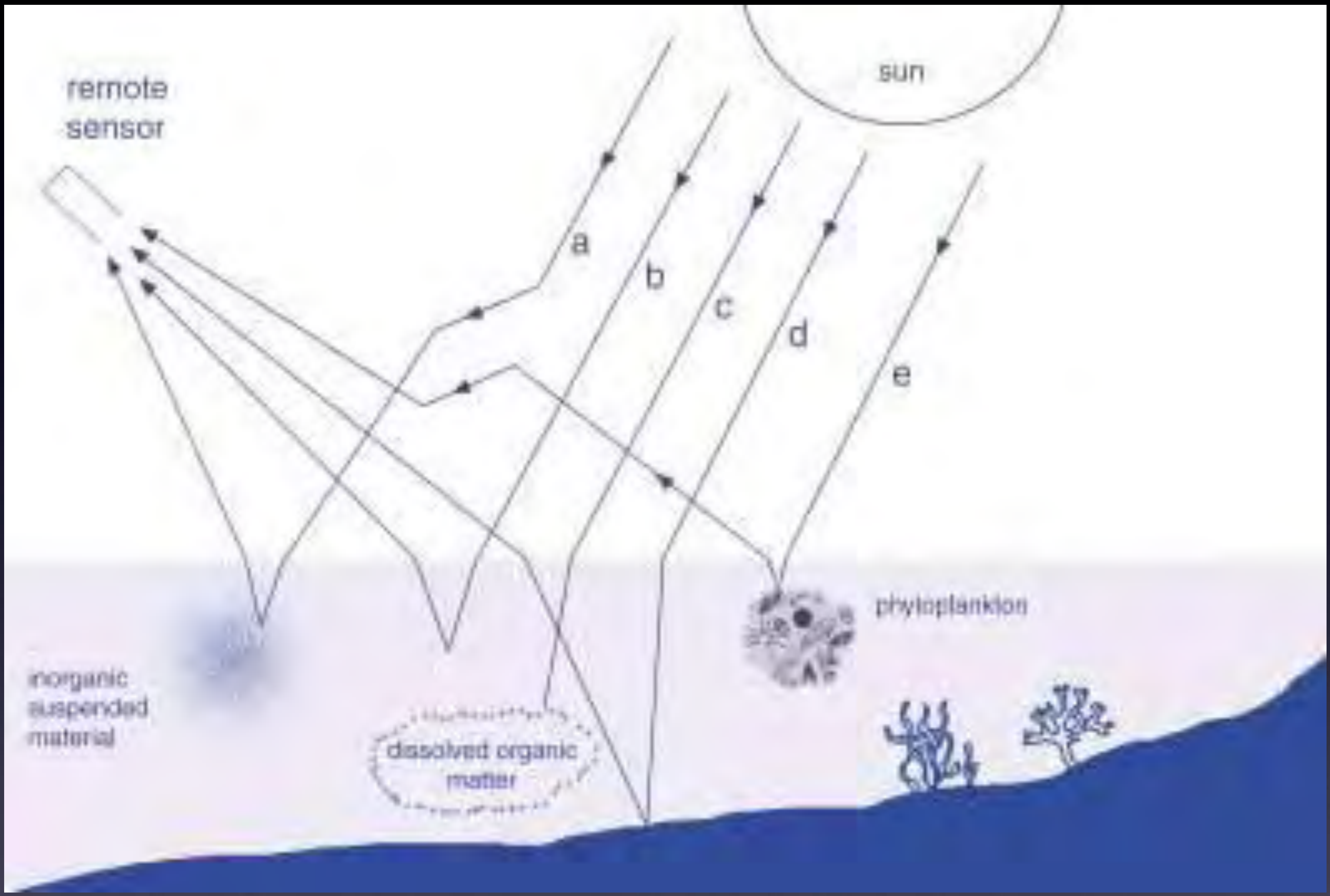
In the coastal ocean or near the coastal margins, materials derived from the land also affect the water color. These are called "Case 2" waters.

In Case 2 waters, there are at least three substances, varying independently, that affect the water color. Specifically, the three substances are:

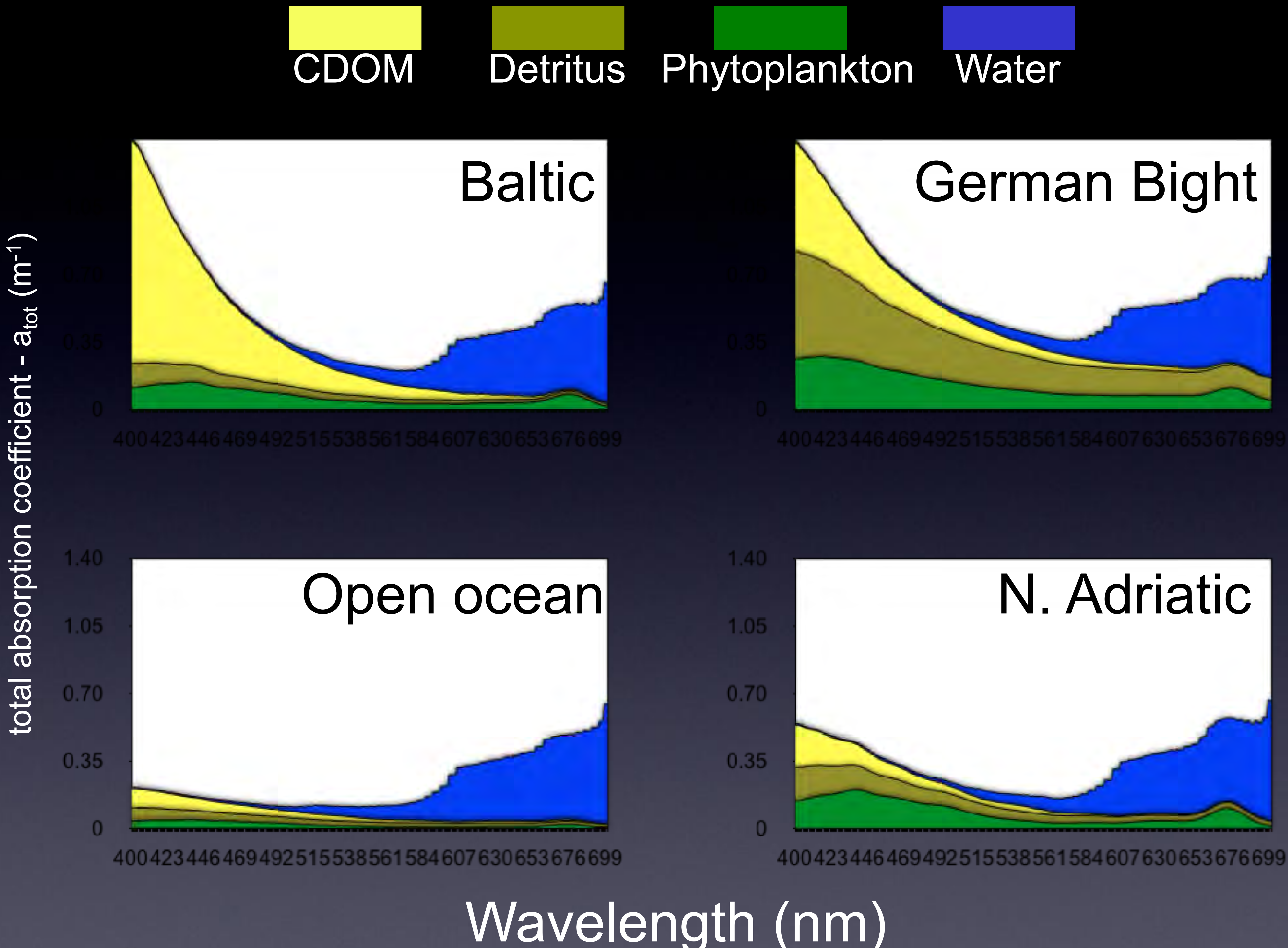
phytoplankton chlorophyll (and other pigments)

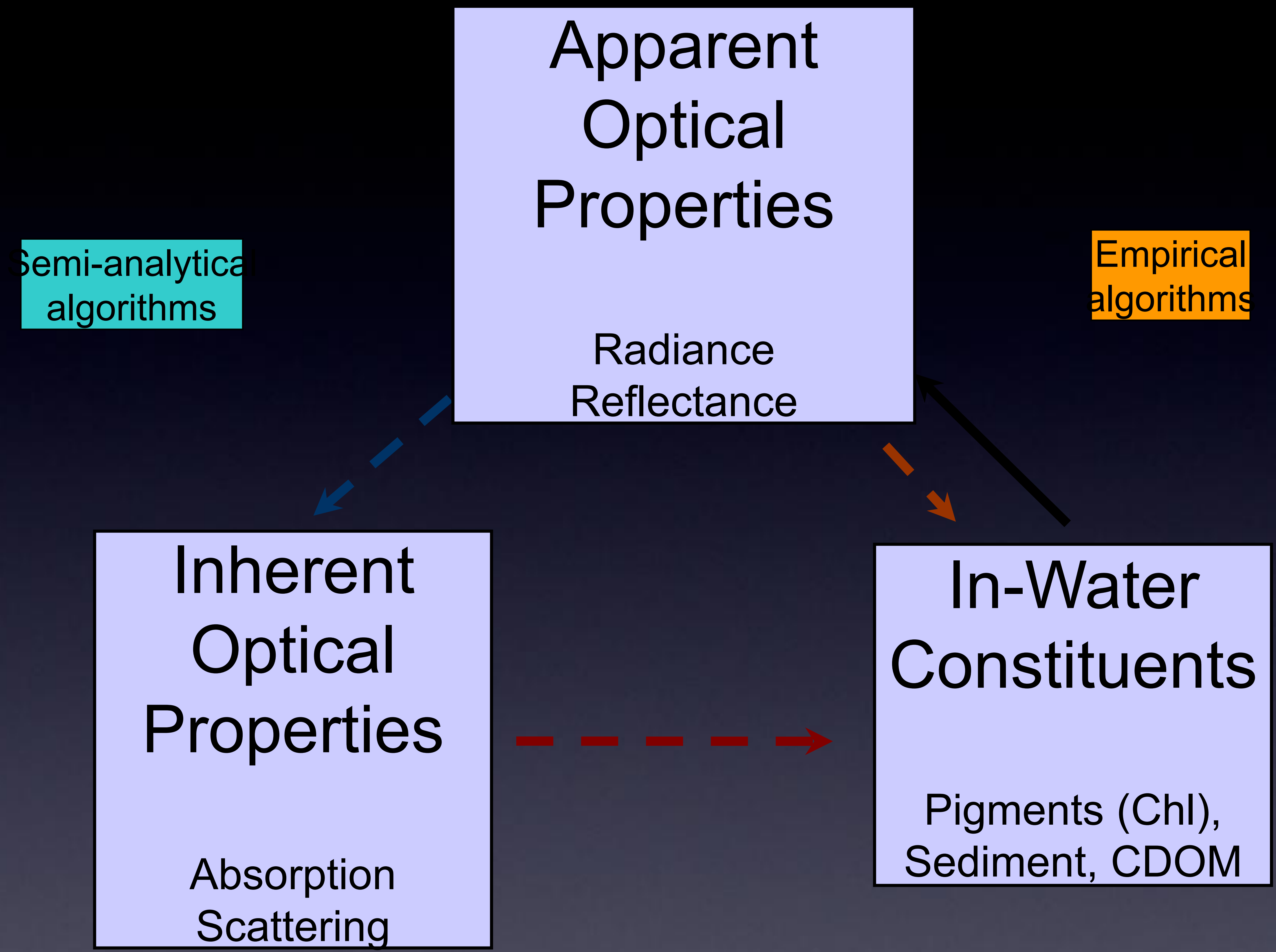
colored dissolved organic matter (CDOM)

non-living particles (sediments and organic detritus)

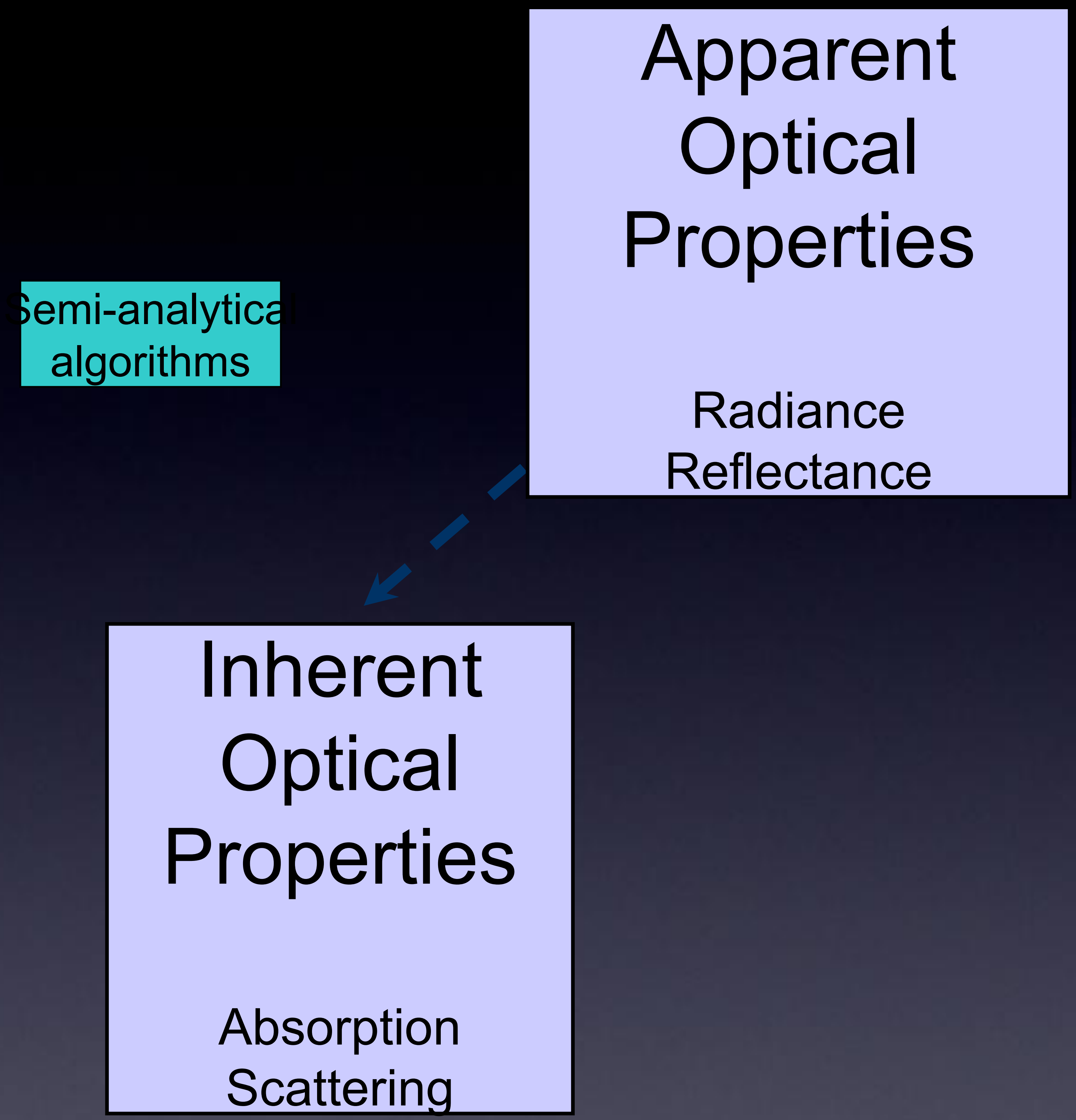


remote sensor
sun
a
b
c
d
e
phytoplankton
inorganic suspended material
dissolved organic matter

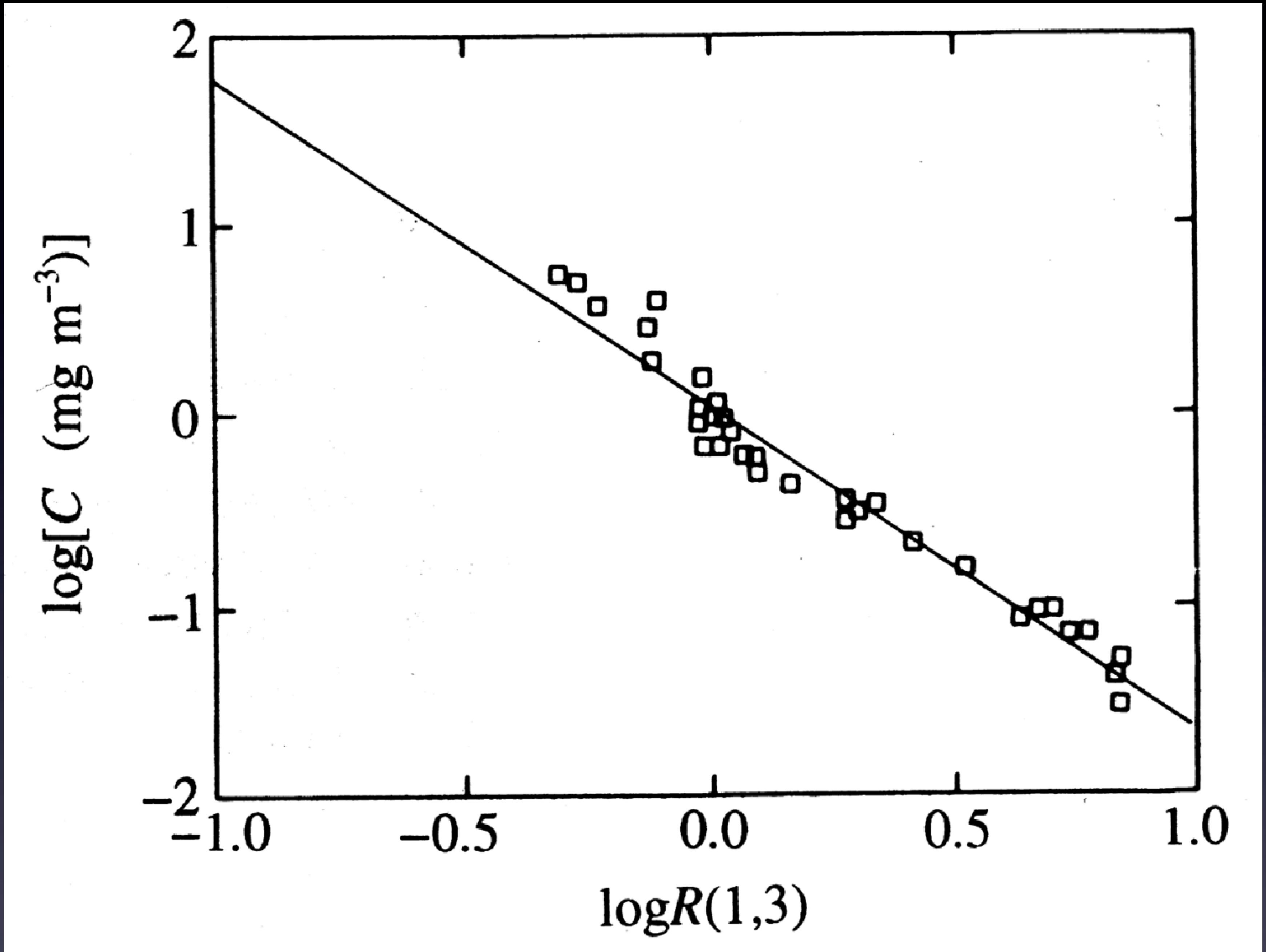

CDOM
Detritus
Phytoplankton
Water
Baltic
German Bight
Open ocean
N. Adriatic
total absorption coefficient - a_{tot} (m^{-1})
1.40
1.05
0.70
0.35
0
400 423 446 469 492 515 538 561 584 607 630 653 676 699
Wavelength (nm)



Apparent Optical Properties
Radiance
Reflectance
Semi-analytical algorithms
Empirical algorithms
Inherent Optical Properties
Absorption
Scattering
In-Water Constituents
Pigments (Chl), Sediment, CDOM


Semi-analytical algorithms
Apparent Optical Properties
Radiance
Reflectance
Inherent Optical Properties
Absorption
Scattering

$R(1,3) = L_w(B_1{=}443)/L_w(B_3{=}550)$ vs Chl



Note: only 33 data points were initially available!

Suggests the band-ratio model: $\log_{10}(\text{Chl}) = C_1 + C_2 \log_{10} [L_w(443)/L_w(550)]$
C_1 and C_2 are the model parameters whose values are determined by the data

Apparent Optical Properties

Remote Sensing Reflectance

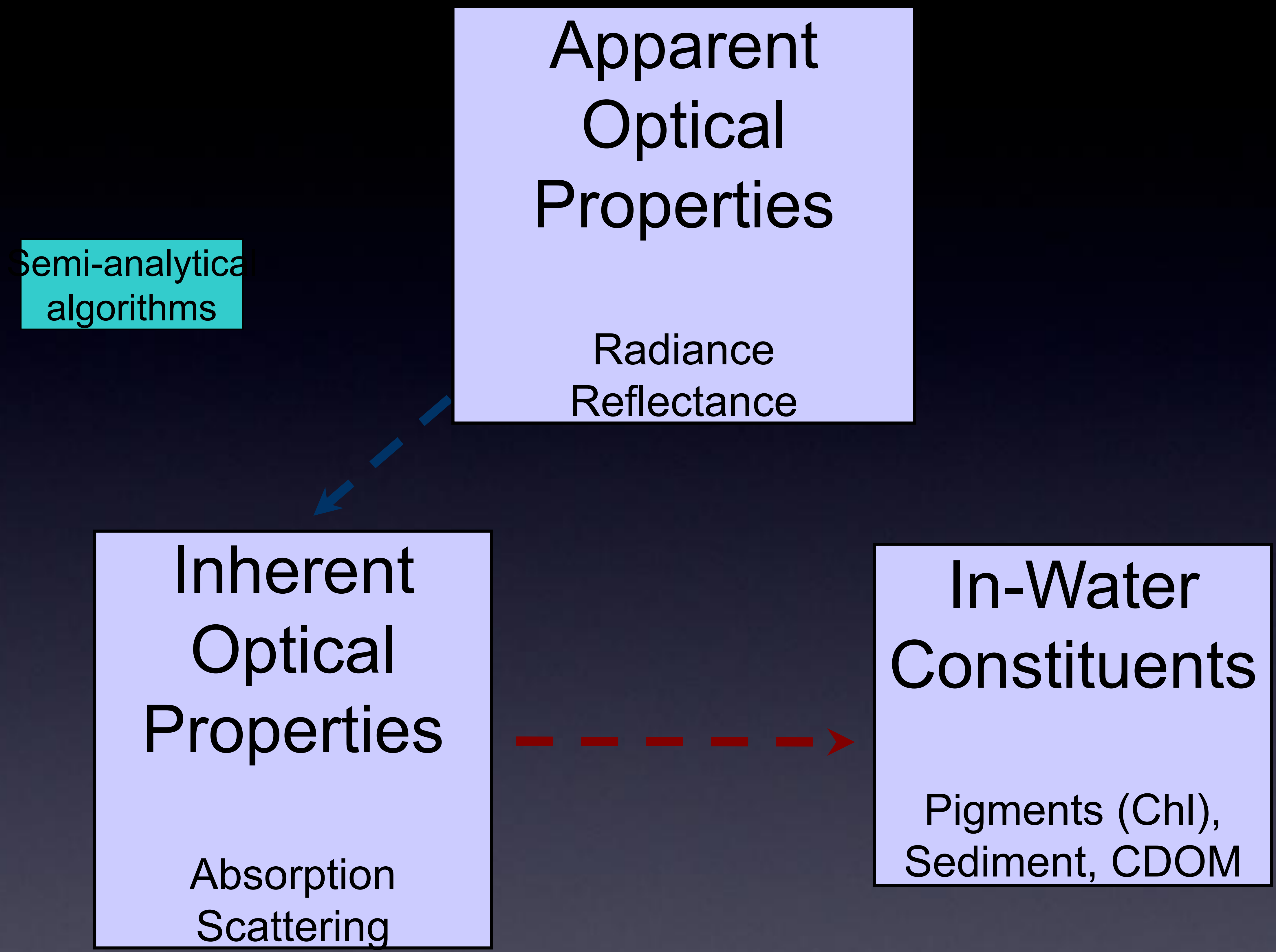
Inherent Optical Properties

Absorption
Backscattering

$$R_{rs}(\lambda) = \left[\frac{f(\theta,\phi)}{Q(\theta,\phi)}\right]\left[\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}\right]$$

Sun/Viewing Geometry

..... and IOPs (a, VSF)



Local parameterization for coastal & inland waters

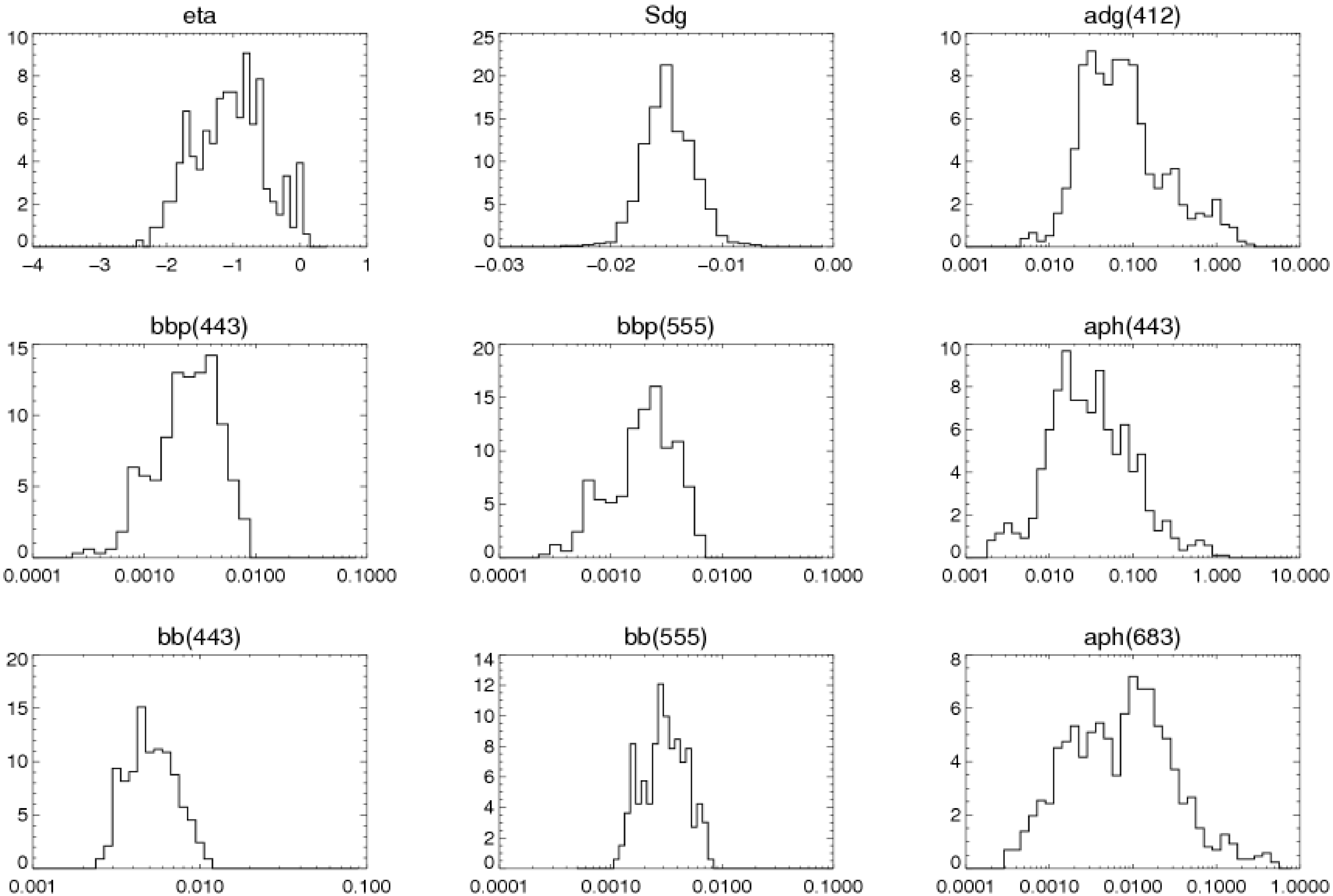
$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda, Chl) + a_d(\lambda, TSS) + a_{cdom}(\lambda, CDOM)$$

$$a(\lambda) = a_w(\lambda) + Ac(\lambda)[Chl]Bc(\lambda) + [a_{cdm}(440)]\ exp(-S(\lambda - 440));$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda, Chl, TSS)$$

$$b_b(\lambda) = b_{bw}(\lambda) + [b_{bp}(555)]\ [555/\lambda]^{Y}$$

frequency distributions of IOP products in NOMAD

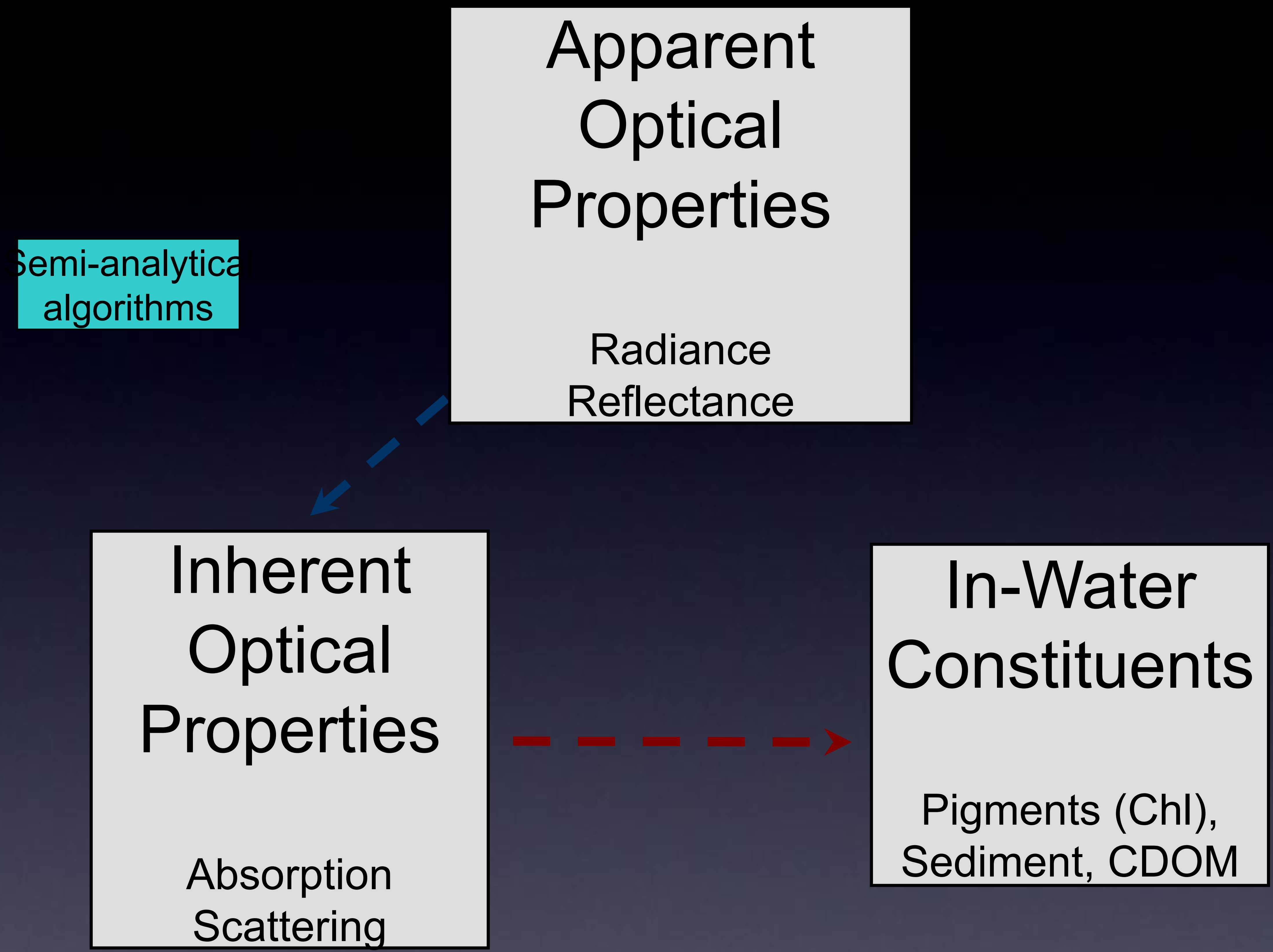


Two step process

- Distinction of different optically active constituents
- Quantification of individual optically active constituents

Section 3

Alternative inversion methods and comparisons


Apparent Optical Properties
Radiance
Reflectance
Semi-analytical algorithms
Inherent Optical Properties
Absorption
Scattering
In-Water Constituents
Pigments (Chl),
Sediment, CDOM

Inversion methods

- Direct inversion/Non linear optimization NLO
- Principal Component Inversion PCI
- Neural Network NN
- Semi-Analytical solution
- Local empirical
- Genetic Algorithms

Availability of methods

- Numerical Recipes
- MATLAB, MATHEMATICA, IDL
- Specific programs: SNNS, NNFit

- All routines can be found as *off the shelf* routines

Synthetic versus in-situ datasets for algorithm training

- Synthetic
  - allow to generate very large datasets for training/parameterisation
  - allow to quantifiably add noise & uncertainty
- In-situ
  - include actual information on the variance and covariance of IOPs and AOPs
  - No assumptions made on bulk IOP AOP relationships (e.g. bi-directonal effects)
  - Independent of specific spectral parameterisation of individual IOP subcomponent

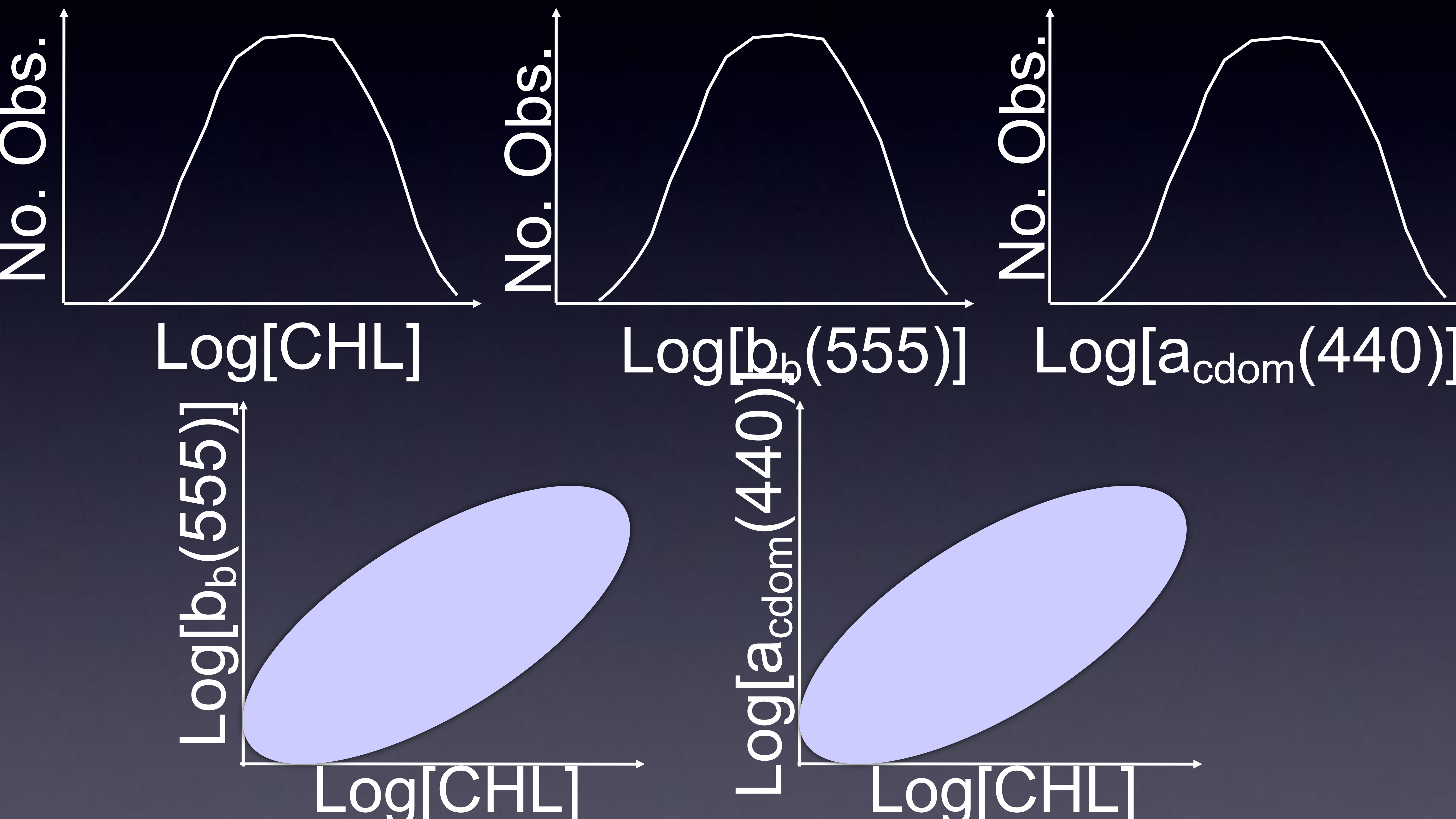
Sensor adopting semi-analytical algorithms

- NLO – SeaWiFS (NASA) trial product
- PCI – MOS (DLR)
- NN – MERIS (ESA), GLI (NASDA)
- Semi-analytical solution MODIS (NASA), planned product for NPP and NPOESS

Critical issues

- An inversion algorithm is only as good as the reflectance (forward) model you use in describing the optical variability of the system/region modelled
- Methods requiring simulated datasets are also highly dependent on the permutation table used i.e. distribution functions of OACs and their covariance

Distribution functions for in-water constituents



OAC covariance - COASTIOOC

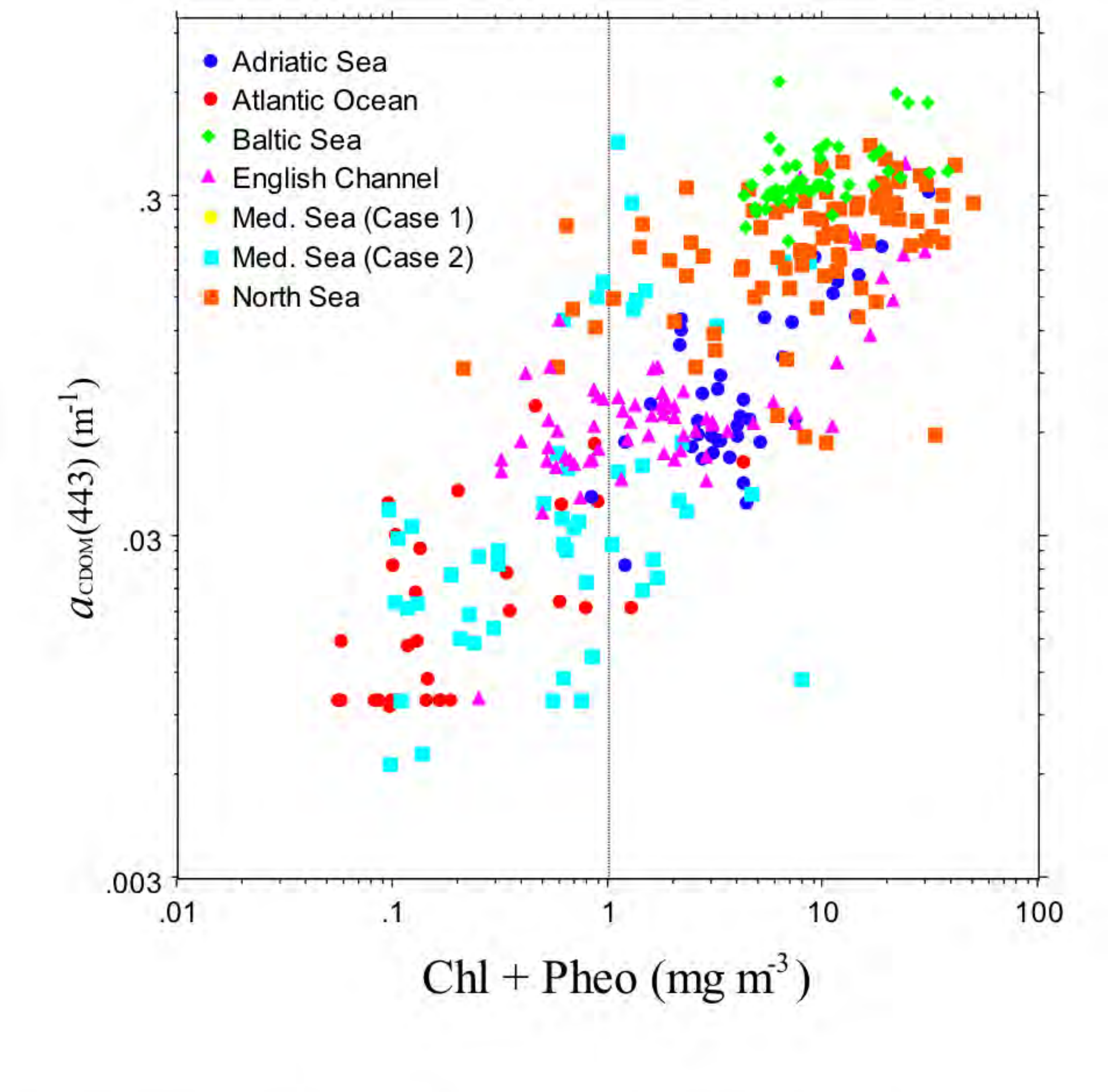


Figure 37. Relationship between Y [represented by $a_{CDOM}(443)$] and Chl + Pheo.

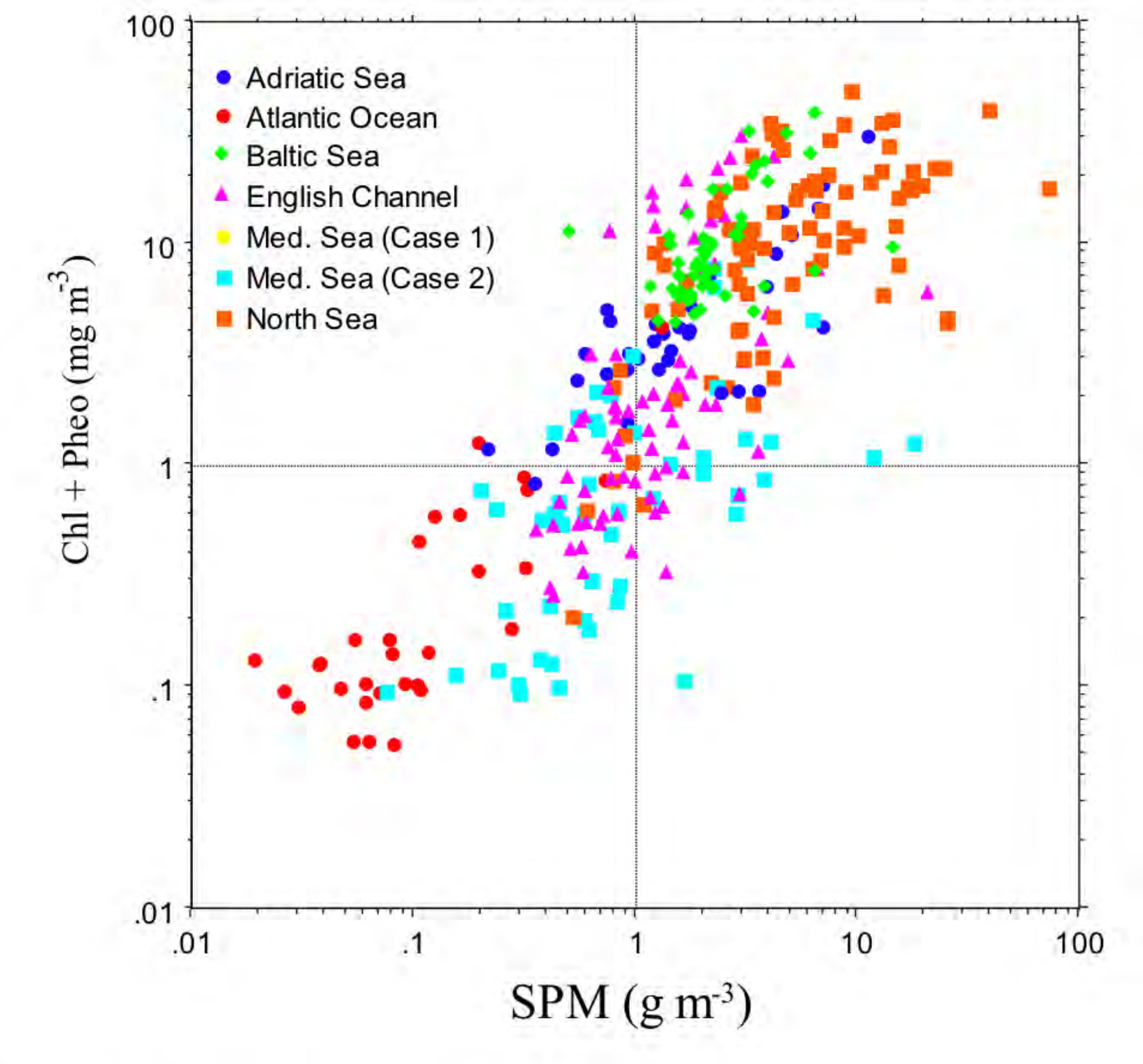


Figure 35. Relationship between Chl + Pheo and SPM.

Babin et. al. (2002)

Non-Linear Optimization

- Basic idea is to minimize the difference between the modelled and measured reflectance until a predefined converge threshold is met

$$\chi^2 = \sum_{\lambda} \left(L_{sat} - L_{\mathrm{mod}} \right)^2$$

- Requires a first guess
- Alternative methods: Levenberg Marquardt, Gauss-Newton, Simplex, differ mainly in the search criteria they use
- Convergence may sometimes be a problem

Apparent Optical Properties

Remote Sensing Reflectance

Inherent Optical Properties

Absorption
Backscattering

$$R_{rs}(\lambda) = \left[\frac{f(\theta,\phi)}{Q(\theta,\phi)}\right]\left[\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}\right]$$

Sun/Viewing Geometry

..... and IOPs (a, VSF)

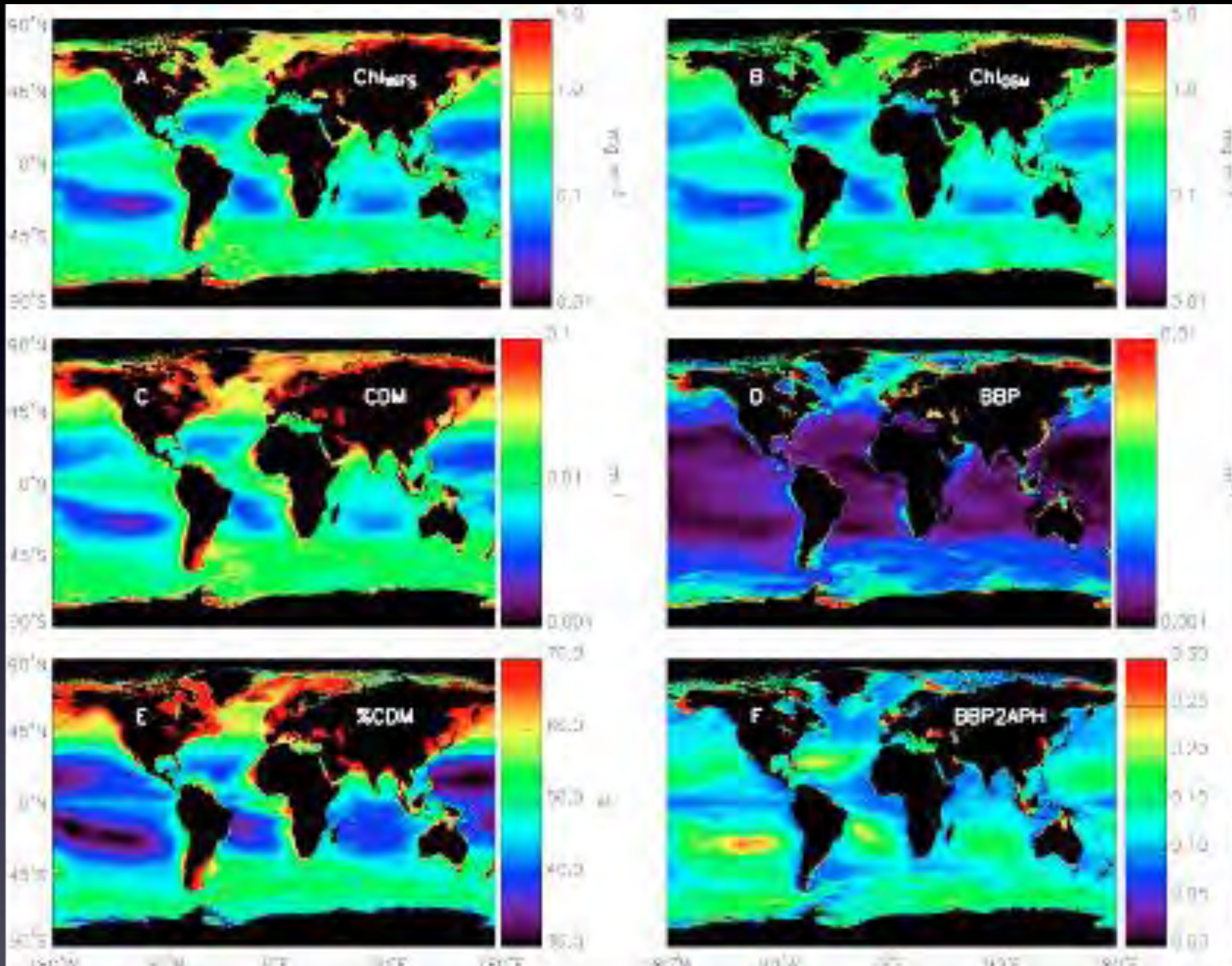
$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda, Chl) + a_d(\lambda, TSS) + a_{cdom}(\lambda, CDOM)$$

$$a(\lambda) = a_w(\lambda) + Ac(\lambda)[Chl]Bc(\lambda) + [a_{cdm}(440)]\, exp(-S(\lambda - 440));$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda, \mathrm{Chl}, \mathrm{TSS})$$

$$b_b(\lambda) = b_{bw}(\lambda) + [b_{bp}(555)]\, [555/\lambda]^{Y}$$

Example NLO application



Siegel et. al. 2005

	NLO	PCI	NN
Advantages	• Direct • Independent of simulated dataset • Non-linear		
Disadvantages	• Convergence • CPU intensive • Initial guess		

Principal Component Inversion

- Principal Component Analysis allows you to extract the most significant information from large multivariate datasets
- Determines a linear transformations of the data to a smaller number of parameters essentially reproducing all of the information of the original data.
- Linear assumption
- Separates signal from noise

Step 1: Forward Model

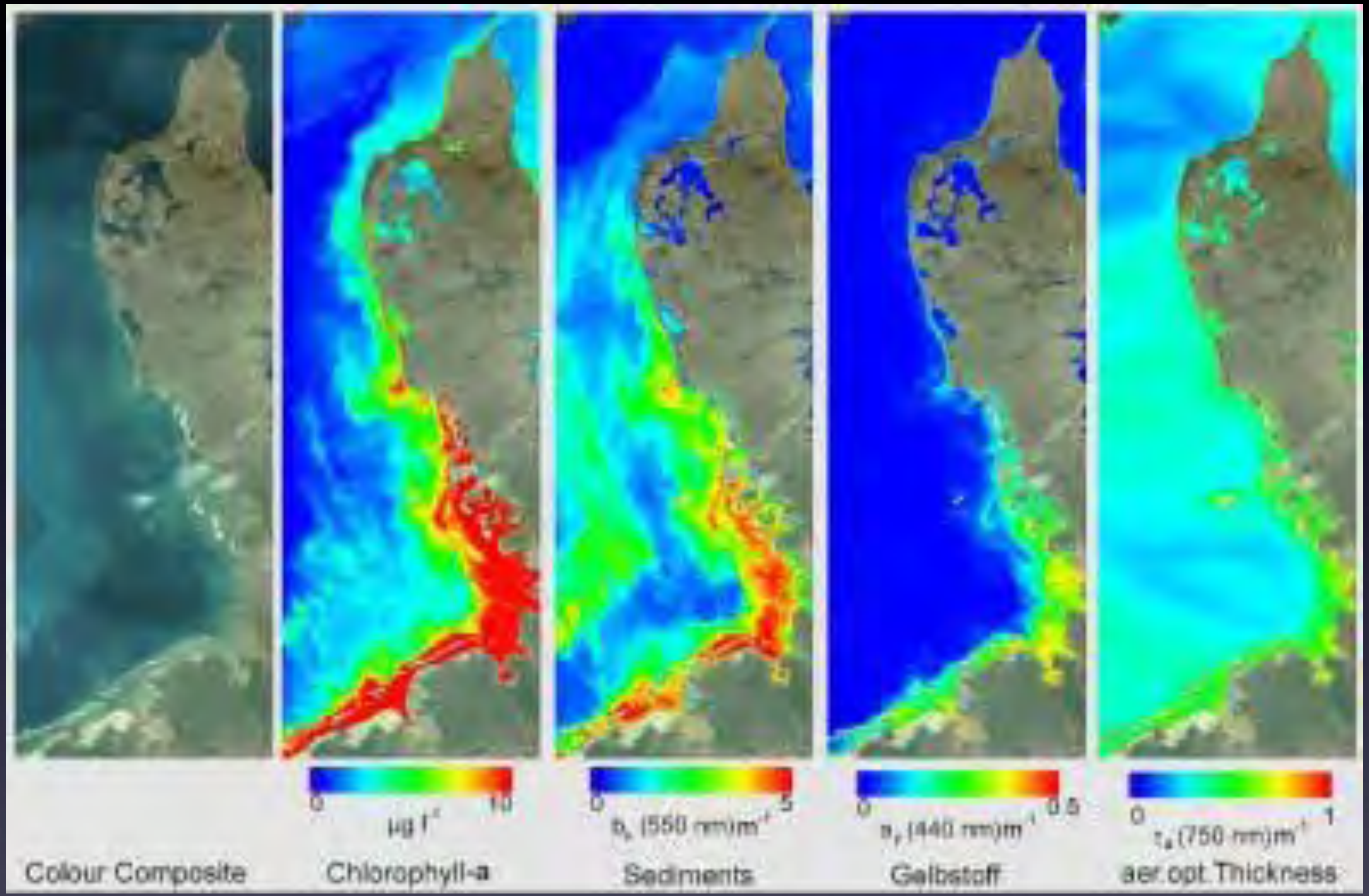
Step 2: Principal Component Analysis

Step 3: Intrinsic Dimensionality

Step 4: Calculation of factor scores

Step 5: Determination of local fit coefficients

Example PCI application



	NLO	PCI	
Advantages	•Direct •Independent of simulated dataset •Non-linear	•Computationally cheap •Always convergence •Separates signal from noise	
Disadvantages	•Convergence •CPU intensive •Initial guess	•Dependent of simulated dataset •Linear assumption	

Neural Network Inversion

- Requires simulated dataset - ~10000 spectra
- Simulated dataset split 2/3 training 1/3 testing
- Training results in “weights file”
- Capable of describing highly non-linear system
- CPU time – high for training v. low for running

Pre-processing

- Log10 transform
- Z-score normalisation

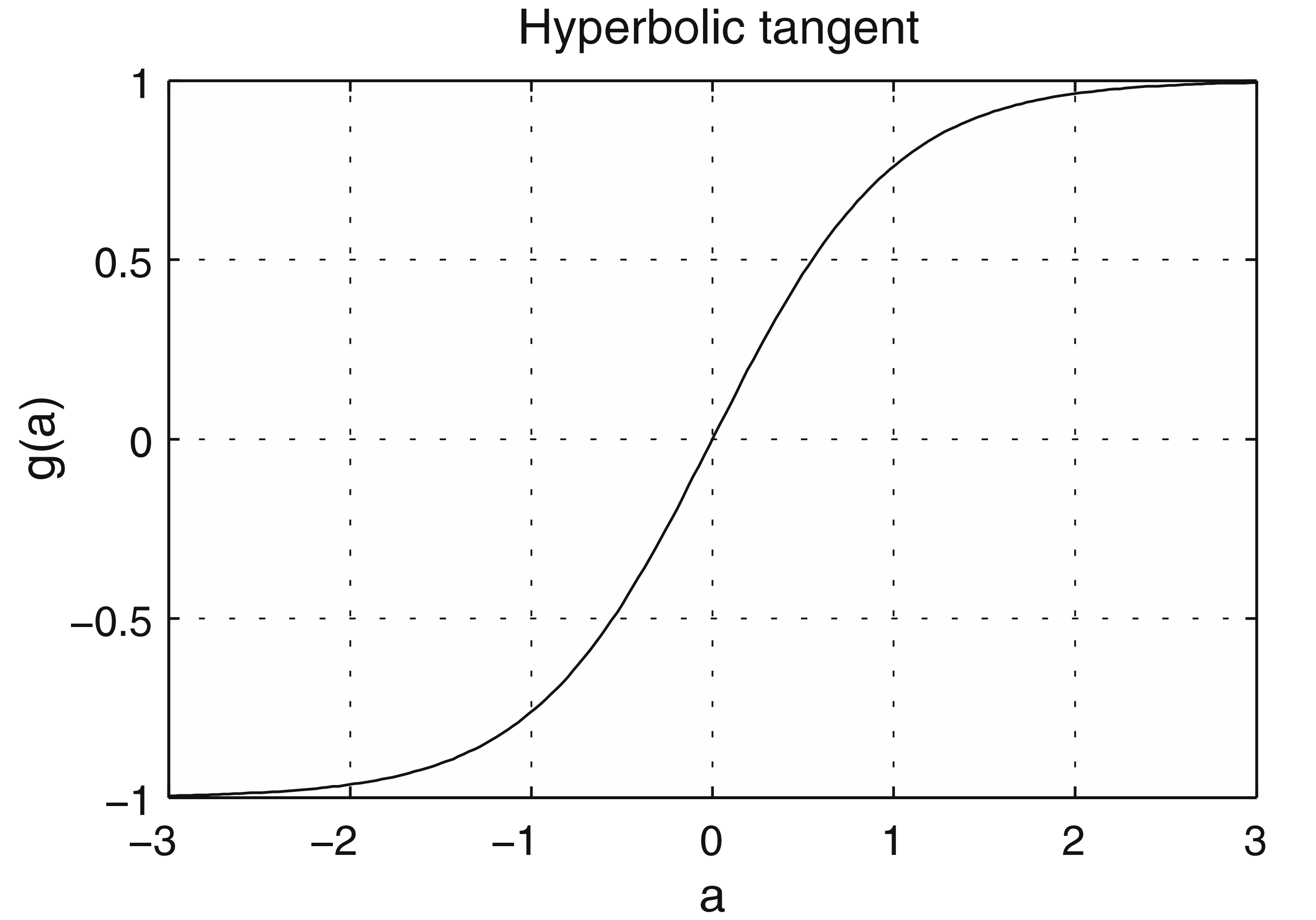
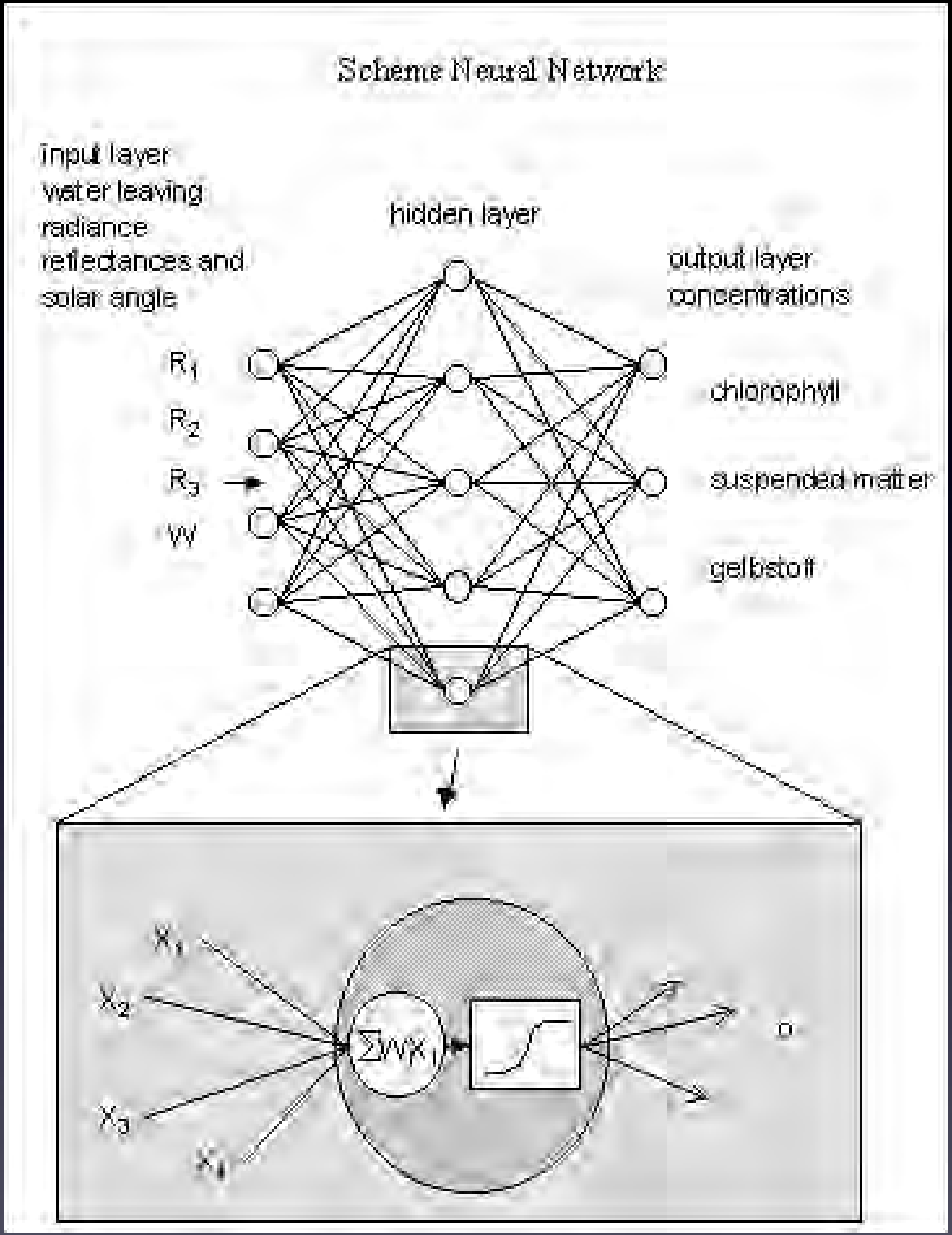


Fig. 1. The hyperbolic tangent activation function.

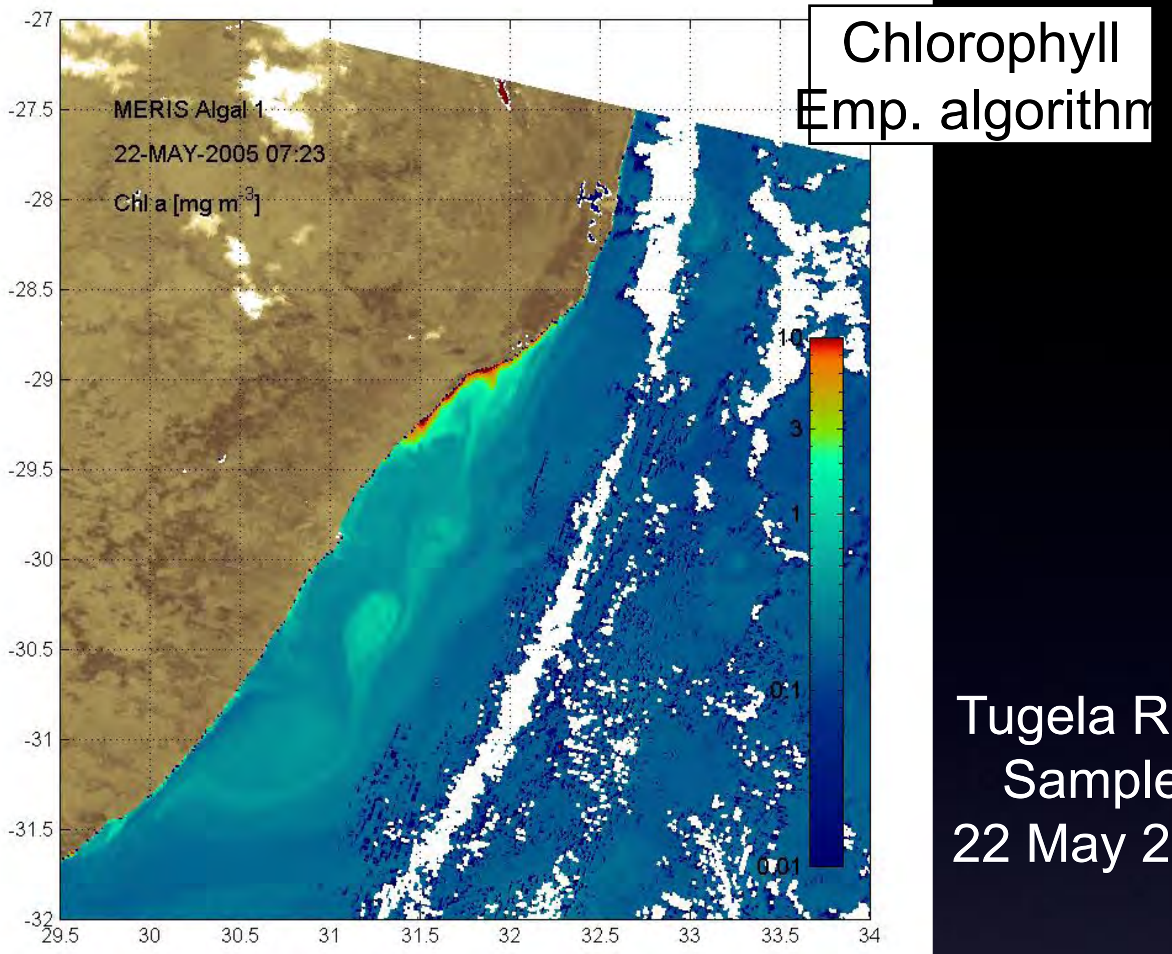

Scheme Neural Network
input layer
water leaving
radiance
reflectances and
solar angle
hidden layer
output layer
concentrations
R1
R2
R3
W
chlorophyll
suspended matter
gelbstoff
X1
X2
X3
X4
ΣWiXi
o

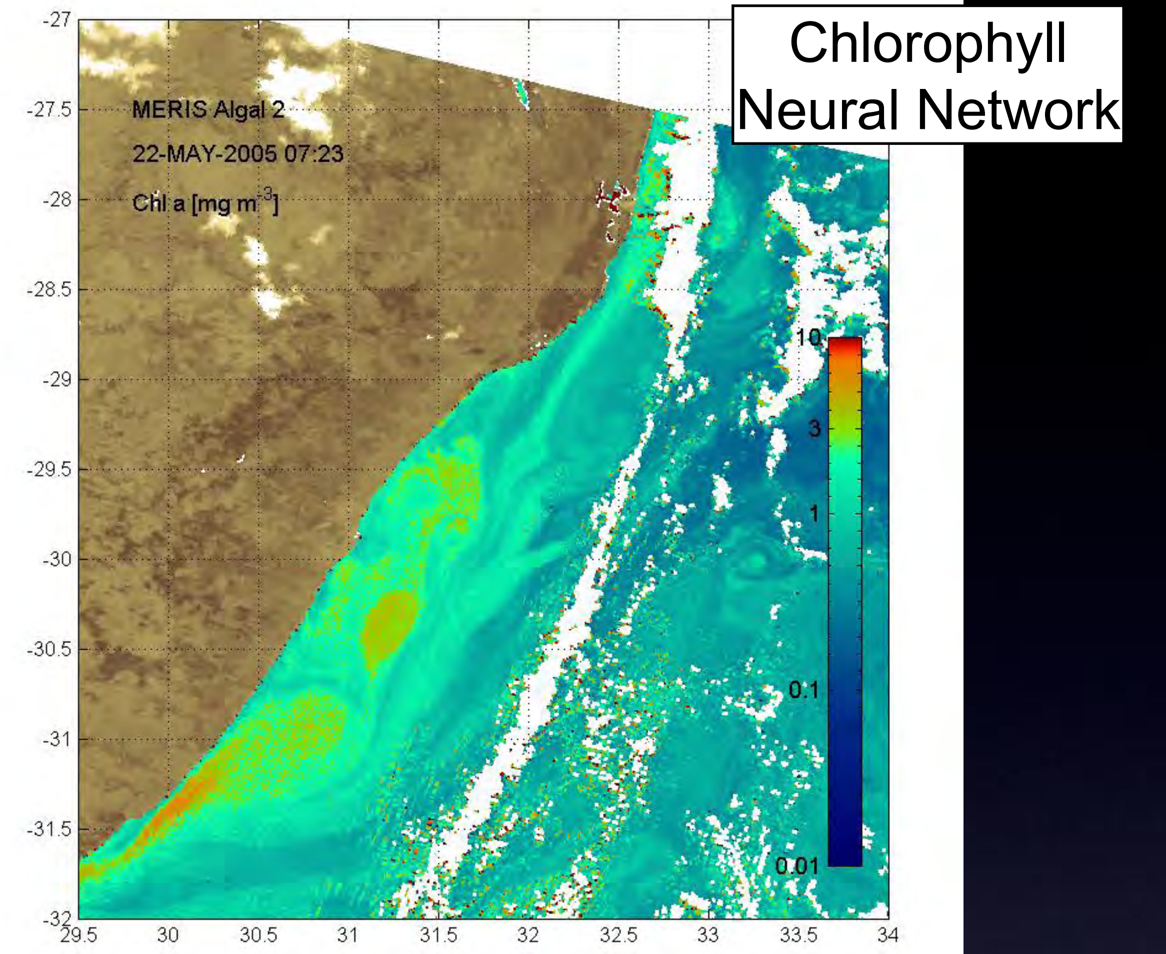
Things to Note

NN use the training data to determine a set of weights so that the given input produced the desired output. After training, we hope (in more complex networks) that new inputs (not in the training data set) will also produce correct outputs.

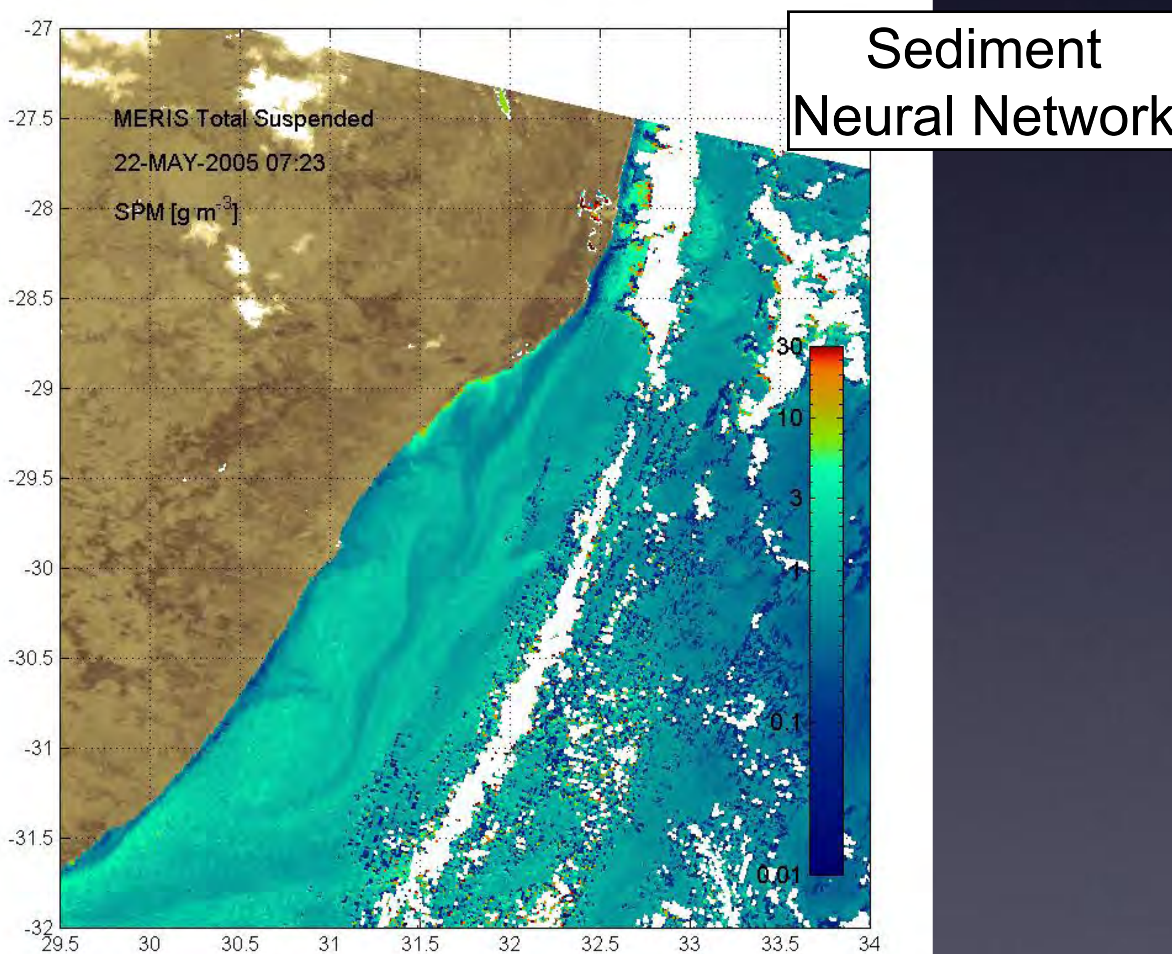
The “knowledge” or “memory” of a neural network is contained in the weights.

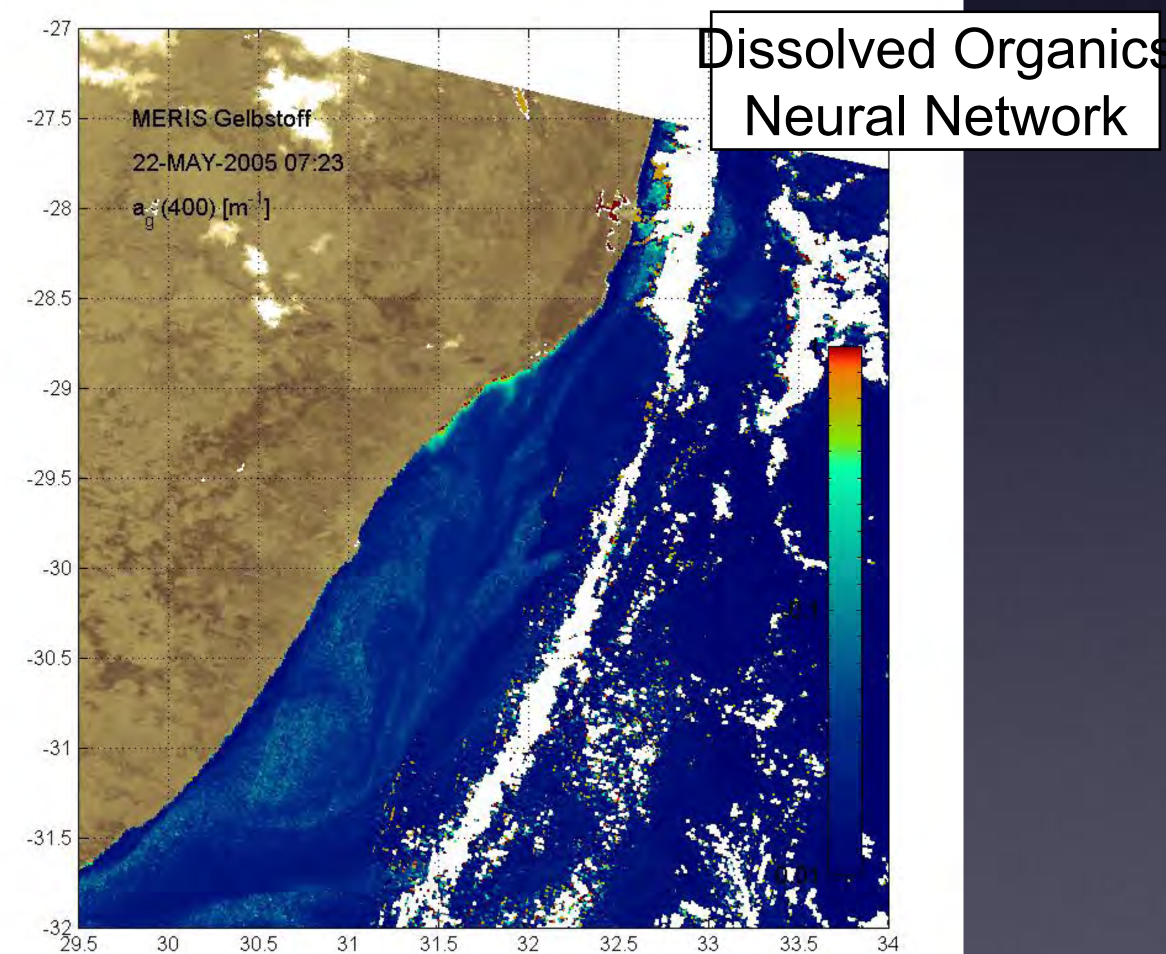
In a more complicated situation, you must balance having enough neurons to capture the science, but not so many that the network learns the noise in the training data.





Tugela River
Samples
22 May 2005





	NLO	PCI	NN
Advantages	•Direct •Independent of simulated dataset •Non-linear	•Computationally cheap •Always convergence •Separates signal from noise	•Computationally cheap •Always converges •Non-linear
Disadvantages	•Convergence •CPU intensive •Initial guess	•Dependent of simulated dataset •Linear assumption	•Dependent of simulated dataset •Slow Training •Choice of architecture

Which is the best?

- Depends on your application, water type
- My suggestion:
  - spend time defining a good reflectance model
  - Characterize distribution functions and covariance of OACs
  - You can always try different inversion method to see which suits your application

Implementing Algorithms – an OCR Agency Perspective

(B. Franz, NASA –pers. com.)

- Require that the algorithm has been validated (at some level) using satellite inputs and field measurement (ideally published)
- Require that the algorithm or product is something of interest to the broader community (requests, forum etc)
- Implement in the NASA processing code: this is more likely to happen if the algorithm has already been implemented in l2gen by the developer… At this stage we also need to be sure that quality screening is adequate (warning and failure conditions trapped).
- Distribute in SeaDAS: this provides an opportunity for the community to test a new algorithm or product, and to see the details of implementation. We may receive feedback through the ocean color forum.

Implementing Algorithms (2)

- Produce global test products: this is a primary function we perform for algorithm development, to take something that has only been done on small scales and show how it performs on the global scale. We may also do global match-up analyses if this is a product for which we have many field measurements. There may be some iteration with the algorithm developer…At this stage, the algorithm may fail when confronted with the full range of radiant path geometries and water properties, or it may be impractical for global application due to resource requirements, and thus we stop.
- Reprocess and distribute global, life-of-mission Level-3 products for evaluation: for derived products, this is typically done by processing Level-3 Rrs to Level-3 products, which we can do quickly with little resource and no impact to standard products... If they are found to be useful, then we may consider step 7.
- Incorporate the product or algorithm as part of standard Level-2 and Level-3 production, at the next full mission reprocessing.

Day 2

IOCCG WG 5

- Working group on Ocean-Colour Algorithms (Chaired by ZhongPing Lee)
- Report: IOCCG Report 5 (2006). Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications.

Objectives

- The objectives of the group were to perform algorithm cross comparisons, to make recommendations on specific algorithms and to report on the progress of algorithm development.
- The group assembled a database from in situ measurements and also developed a synthesized dataset based on known relationships, in order to perform algorithm cross-comparisons and evaluations. The synthesized datasets, as well as the software for the various algorithms, are available on the IOCCG webpage.

Terms of Reference

- Synthesize a database of inherent (IOP) and apparent optical properties (AOP), and assemble a database of in situ measurements.
- Perform cross-comparisons and evaluations on existing ocean-colour inversion algorithms.
- Make recommendations on specific algorithms.
- Report on the progress of algorithm development.

Algorithms

Algorithm	Type	Key features
L98	Empirical	Empirical constants; products at 440 nm only
B99	Semi-empirical	Relationships between total absorption coefficients
MM01		Bio-optical models; hyperspectral
Loisel		$K_d(\lambda)$ from $R_{rs}(\lambda)$ empirically
D&S	Neural Network	Neural constants; MERIS only
Lyon	Algebraic *(Linear Matrix Inversion)*	Spectral models for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $b_{bp}(\lambda)$
Boss		Varying spectral shapes for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $b_{bp}(\lambda)$; statistical selection of solution; generates output confidence intervals; applicable to multi- and hyperspectral data
Carder	Algebraic for low absorption waters (*iterative solution*); empirical for other	Spectral models for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $b_{bp}(\lambda)$; empirical coefficients for different properties
QAA	Algebraic	Separate derivations for the total and individual components; spectral models for $a_{dg}(\lambda)$ and $b_{bp}(\lambda)$; retrieve multi- or hyperspectral a_{ph} spectrum
GSM	Spectral optimization	Optimized spectral shapes for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $b_{bp}(\lambda)$; applicable to multi- and hyperspectral data; can use input uncertainties and generates output confidence intervals
SPD		Varying spectral shapes for $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, and $b_{bp}(\lambda)$; applicable to multi- and hyperspectral data

Simulated vs In-situ dataset

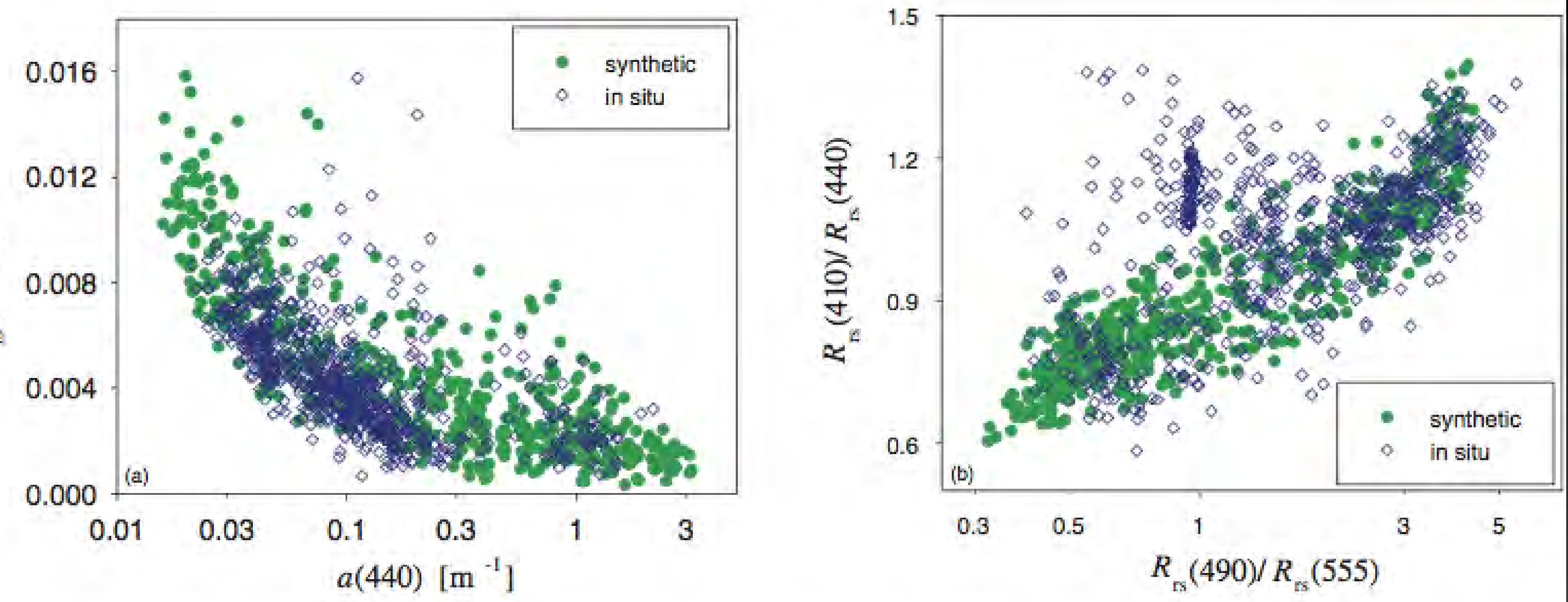
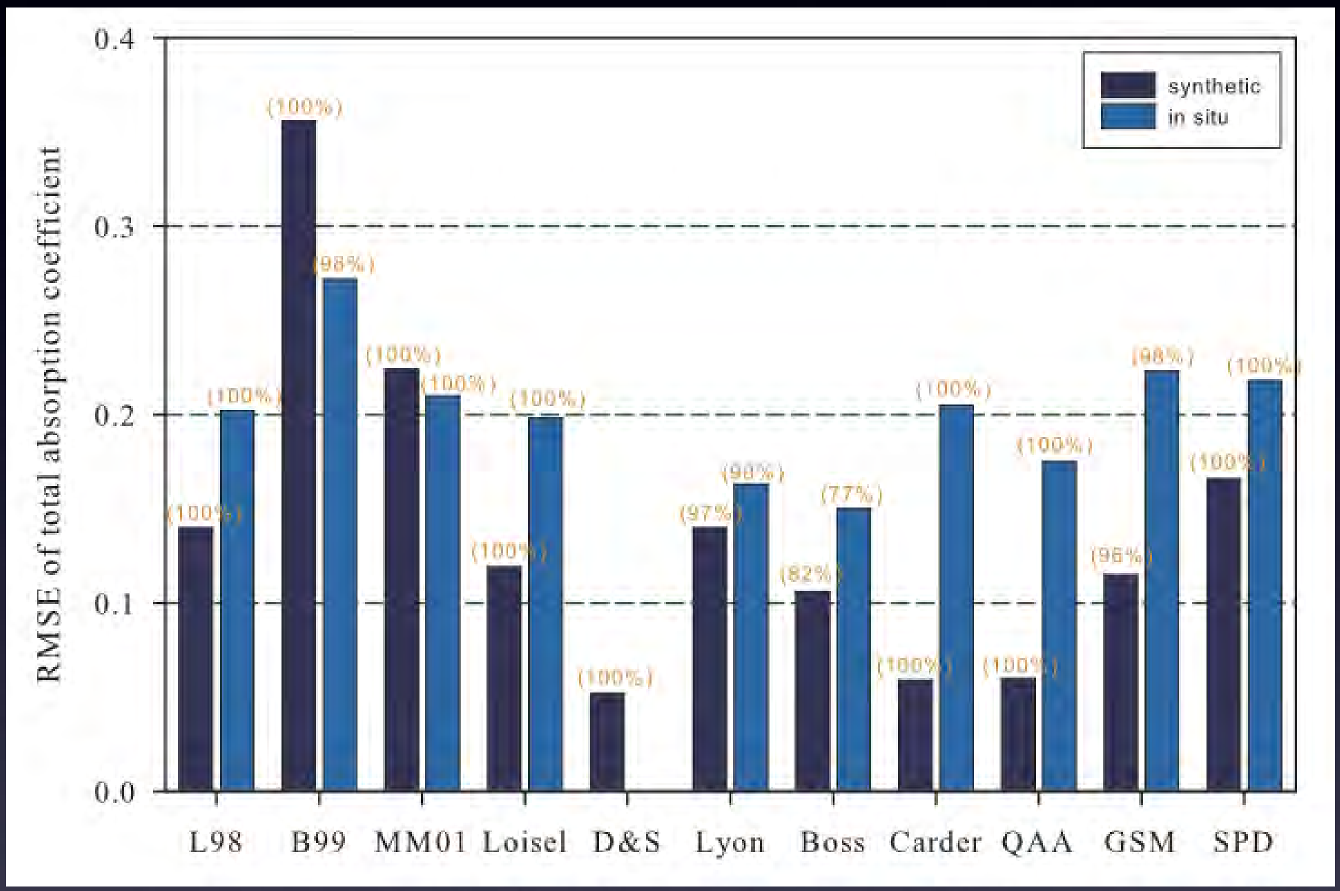
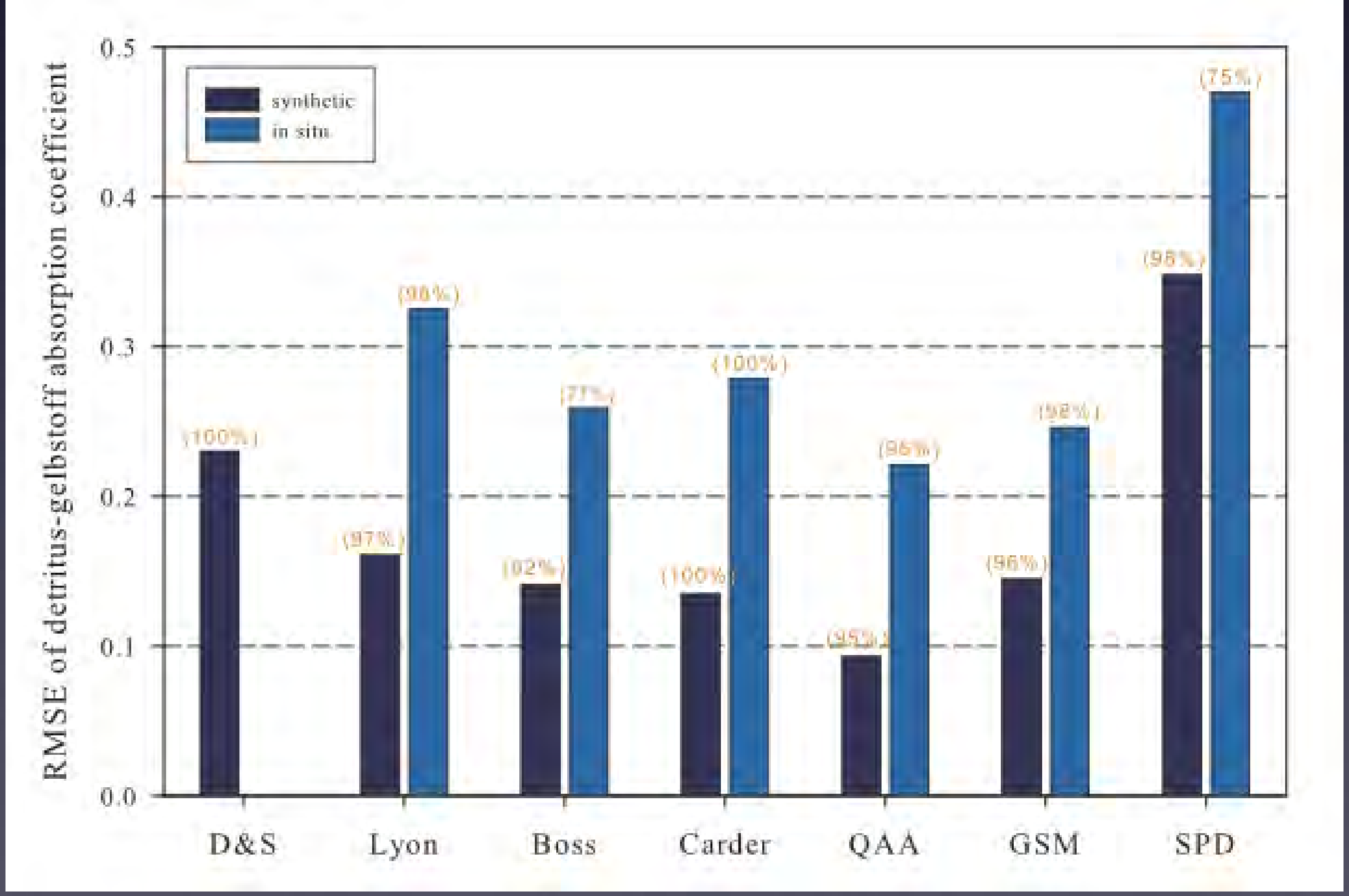


Figure 2.3 Comparison between *in situ* and synthetic data sets. (a) Ranges and variations of $R_{rs}(440)$ and $a(440)$. (b) Ranges and variations of $R_{rs}(410)/R_{rs}(440)$ and $R_{rs}(490)/R_{rs}(555)$.

Bulk vs Constituent retrievals





Conclusions (1)

- In general,the best properties that can be obtained from ocean-colour data, regardless of the algorithm used … are the spectral absorption and backscattering coefficients of the total water volume.
- Using the synthetic data set as a reference, more reliable results are obtained for clearer waters ($a(440) < \sim 0.3\ m^{-1}$)… than more absorbing waters ($a(440) > \sim 0.3\ m^{-1}$).
- When decomposing the total absorption coefficient into the components of phytoplankton and coloured material, less accurate results are anticipated due to overlapping of spectral signals and because the spectral shapes of the components are not constant.
- If the chlorophyll-a concentration (C) is desired from ocean colour, more uncertainties will be introduced because the chlorophyll-specific absorption coefficient is not constant at a given wavelength…

Conclusions (2)

- The robust and stable results of the total absorption and backscattering coefficients from these various algorithms, …these optical properties should be taken as standard products for all ocean-colour satellite missions. …
- Space-based sensors should be equipped with at least one spectral band in the region of 620-640 nm. Such a band is very important for coastal remote sensing ...
- Algorithms based on the fundamentals of hydrological optics are strongly advocated…

Recommendations

- Increased high-quality, co-located measurements of remote-sensing reflectance and IOPs.
- Improved methods to select model parameters such as the spectral shapes of individual IOPs
- Better quantification of uncertainties in derived products
- Improved procedure for atmospheric correction.
- And, finally, enhance and broaden applications

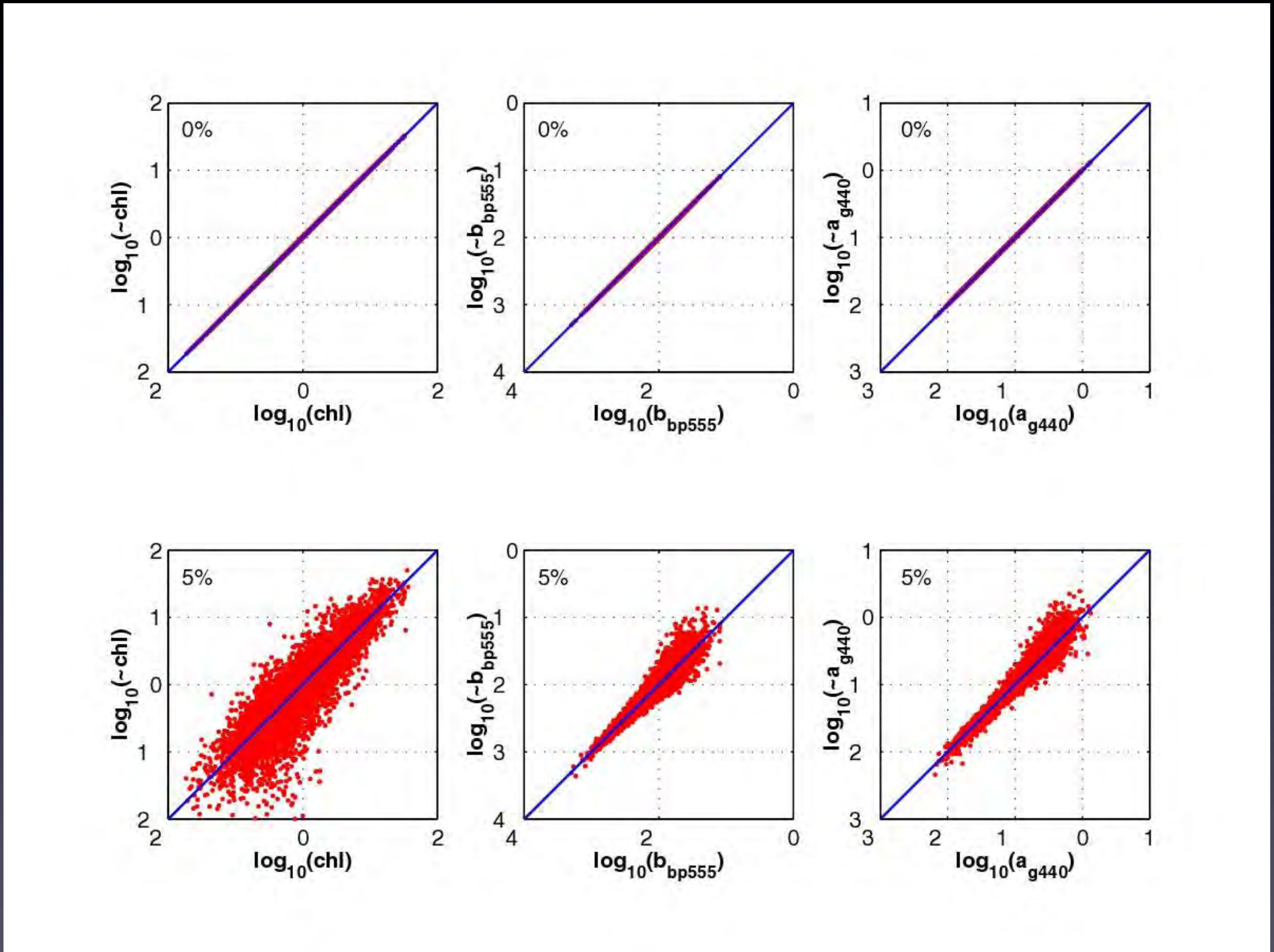
Inversion method inter-comparison

Study undertaken at UNH with OPAL PhD student Hui Feng

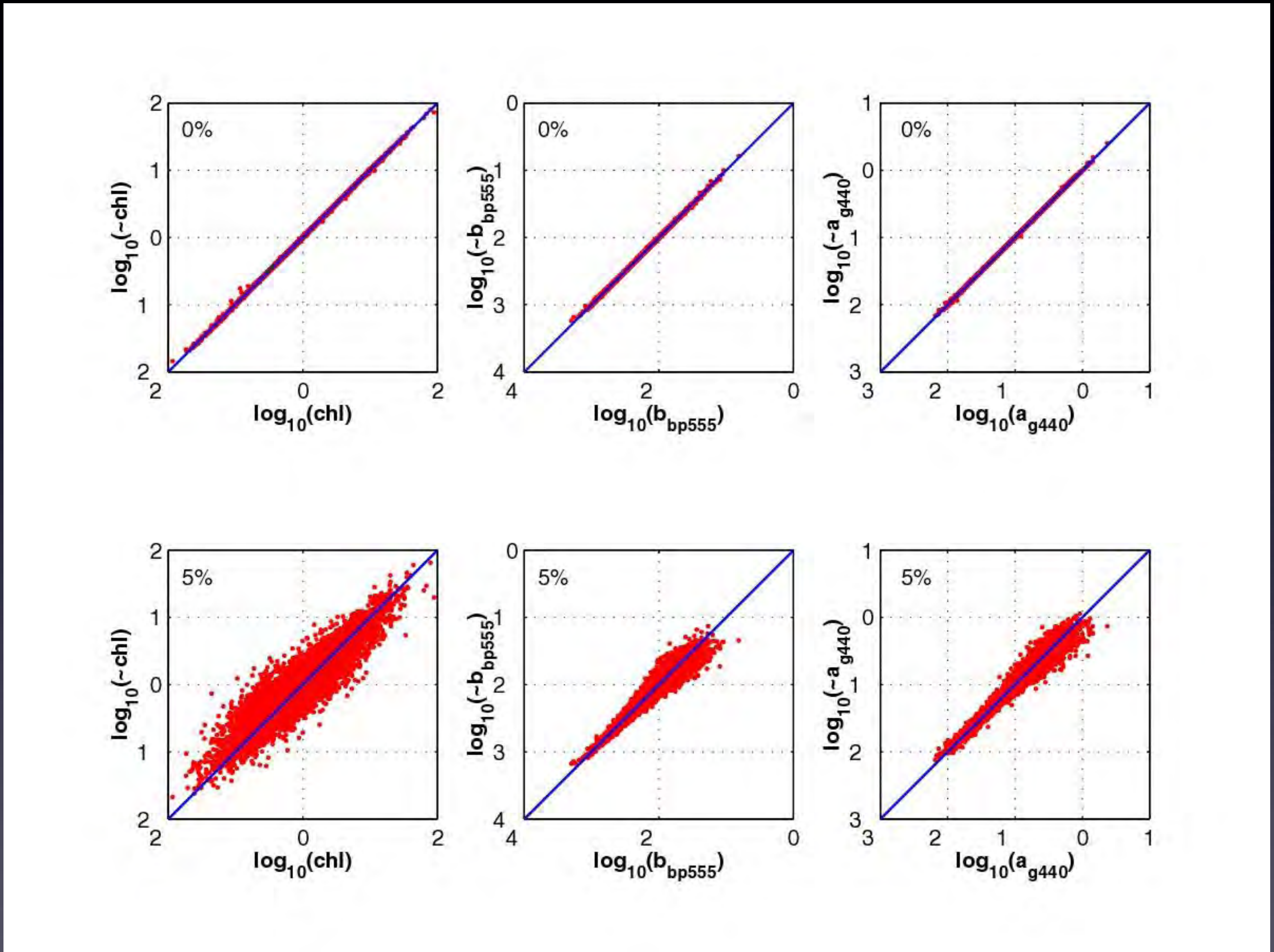
Objectives:

For the different inversion algorithms and optically distinct water classes:
Estimate errors in the retrieved C quantitatively, and
Quantify the effect of the noise-contained ocean color signals on the accuracy in the retrieved C.

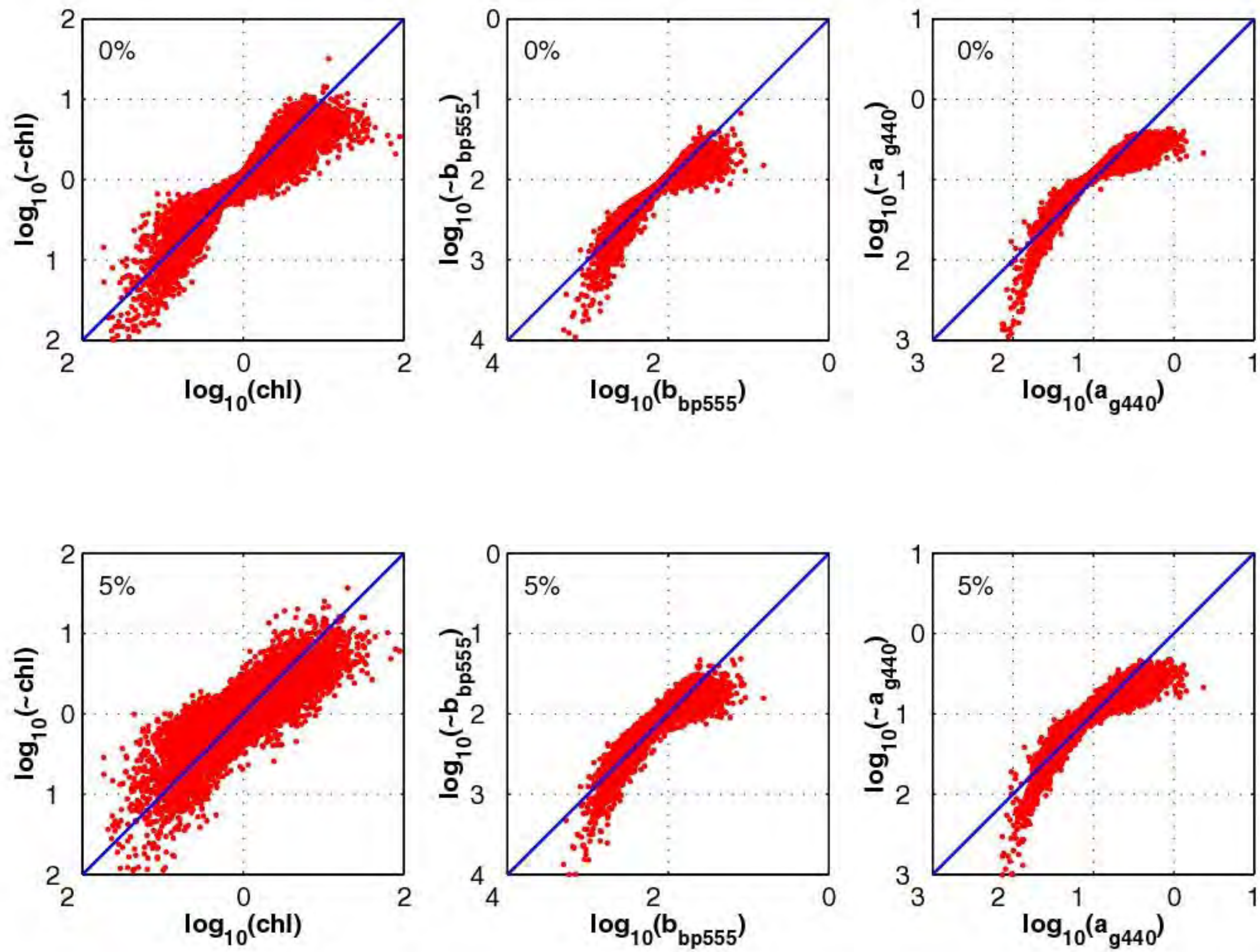
Non-Linear Optimization



Neural Network



Principal Component Inversion



Spectral Unmixing - QAA

Table 1
Steps of deriving phytoplankton absorption spectrum from hyperspectral remote-sensing reflectance

Steps	Property	Derivation
Step 0	r_{rs}	$=R_{rs}/(0.52+1.7R_{rs})$
Step 1	$u(\lambda)=\dfrac{bb(\lambda)}{a(\lambda)+bb(\lambda)}$	$=\dfrac{-0.0895+\sqrt{(0.0895)^2+4g_1r_{rs}(\lambda)}}{20.1247}$
Step 2	$a(\lambda_0)$: $a(555)$ or $a(640)$	$a(555)=0.0596+0.2(a(440)_i-0.01)$, $a(440)_i=\exp(-1.8-1.4\rho+0.2\rho^2)$, $\rho=\ln(r_{rs}(440)/r_{rs}(555))$
Step 3	$bb_p(\lambda_0)$	$=\dfrac{u(\lambda_0)a(\lambda_0)}{1-u(\lambda_0)}-bb_w(\lambda_0)$
Step 4	Y	$=2.2(1-1.2e^{-0.9r_{rs}(440)/r_{rs}(555)})$
Step 5	$bb_p(\lambda)$	$=bb_p(\lambda_0)\left(\dfrac{\lambda_0}{\lambda}\right)^Y$
Step 6	$a(\lambda)$	$=\dfrac{(1-u(\lambda))(bb_w(\lambda)+bb_p(\lambda))}{u(\lambda)}$
Step 7	$\zeta=a_\phi(410)/a_\phi(440)$	$=0.71+\dfrac{0.06}{0.8+r_{rs}(440)/r_{rs}(555)}$
Step 8	$\xi=a_g(410)/a_g(440)$	$=e^{-S(410-440)}$
Step 9	$a_g(440)$	$=\dfrac{(a(410)-\zeta a(440))}{\xi-\zeta}-\dfrac{(a_w(410)-\zeta a_w(440))}{\xi-\zeta}$
Step 10	$a_\phi(\lambda)$	$=a(\lambda)-a_g(440)e^{-S(\lambda-440)}-a_w(\lambda)$

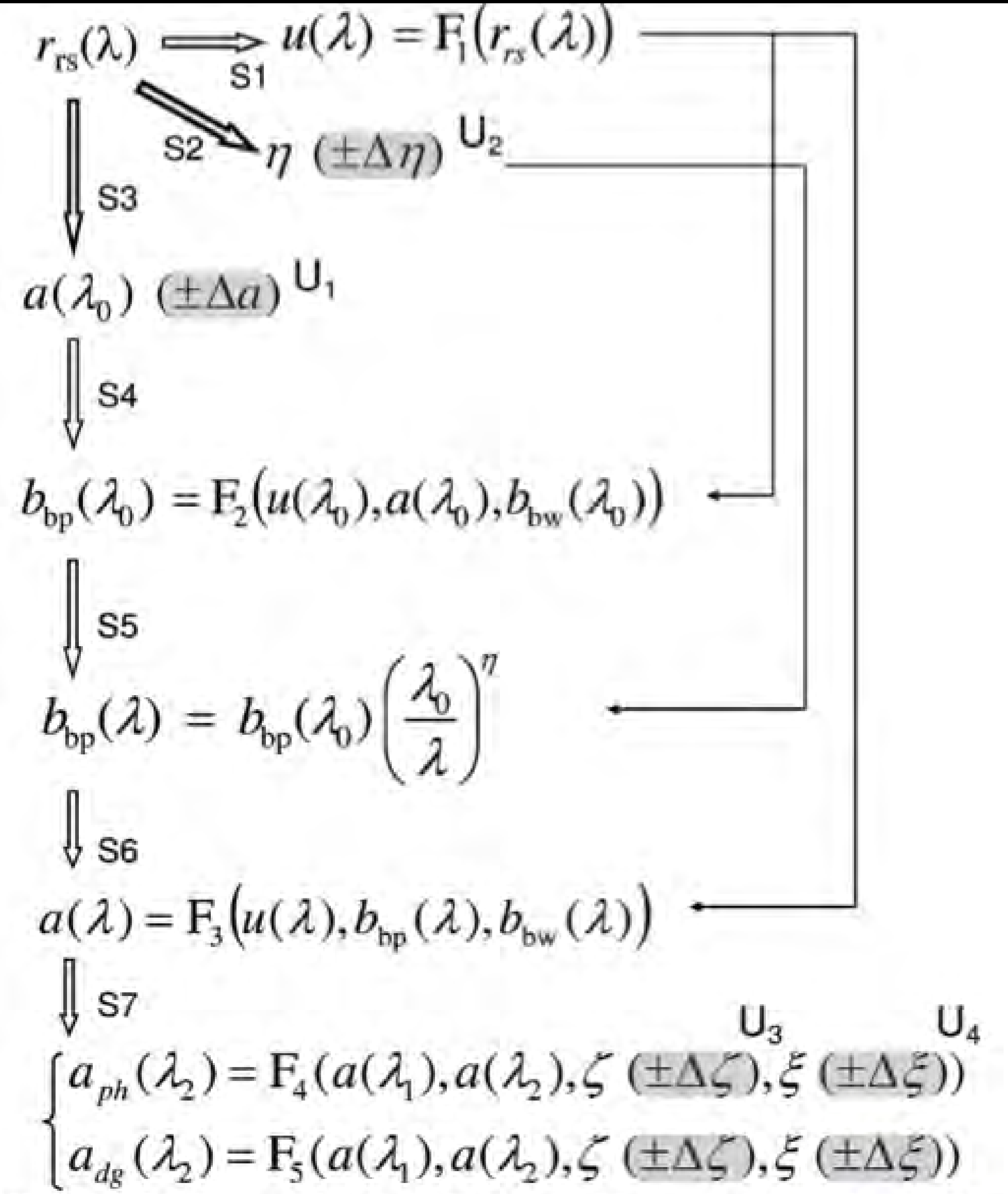


Fig. 1. Schematic chart to show variables and steps (S1–S7) involved in the QAA procedure, redrawn from Lee *et al.* [23]. Variables with uncertainties (U_1–U_4) discussed in this study are highlighted with gray, while all others assumed error free.

construction (& deconstruction) of an SAA …

$$R_{rs} \approx func\left(\frac{b_b}{a+b_b}\right)$$

satellite provides $R_{rs}(\lambda)$

$a\,(\lambda)$ and $b_b\,(\lambda)$ are desired products

1

Spectral Optimization:

* define shape functions for (e.g.) $b_{bp}(\lambda)$, $a_{dg}(\lambda)$, $a_{ph}(\lambda)$

* solution via L-M, matrix inversion, etc.

* ex: RP95, HL96, GSM

2

Spectral Deconvolution:

* partially define shape functions for $b_{bp}(\lambda)$, $a_{dg}(\lambda)$

* piece-wise solution: $b_{bp}(\lambda)$, then $a(\lambda)$, then $a_{dg}(\lambda) + a_{ph}(\lambda)$

* ex: QAA, PML, NIWA

3

Bulk Inversion:

* no predefined shapes

* piece-wise solution: $b_{bp}(\lambda)$, then $a(\lambda)$, via (empirical) $K_d\,(\lambda)$ via RTE

* ex: LS00

consensus to refine spectral optimization to initiate process …

$$R_{rs} \approx func\left(\frac{b_b}{a+b_b}\right)$$

Spectral Optimization:

* define shape functions for (e.g.) $b_{bp}(\lambda)$, $a_{dg}(\lambda)$, $a_{ph}(\lambda)$

* optimization via L-M

our STARTING point:

* dynamic bbp retrieval

* dynamic aph spectral model

* IOP-based f/Q tables

* Raman scattering

* fluorescence

* T/S dependence on aw & bbw

* optical water class parameterization

* uncertainties & propagation of error

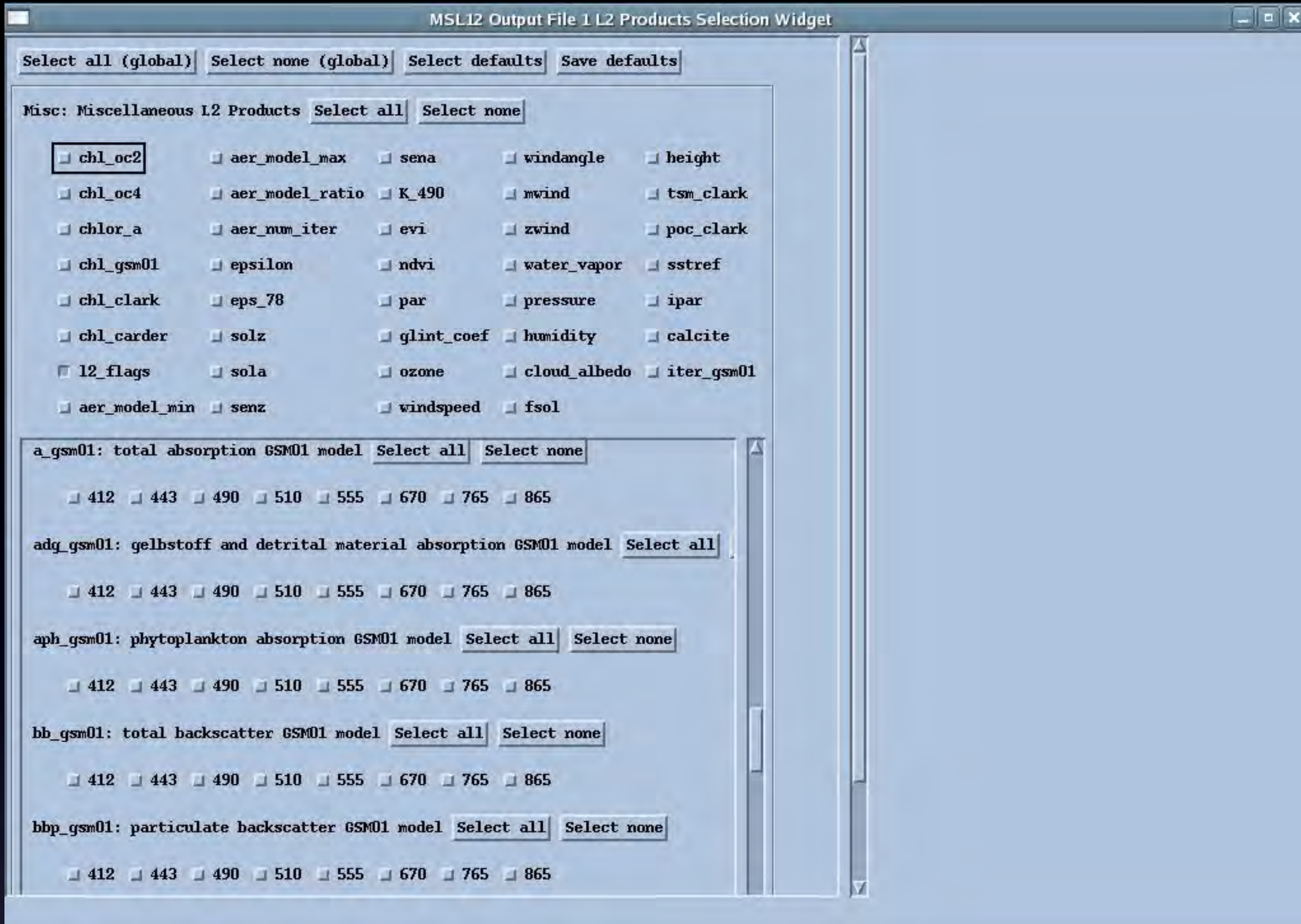
metrics defined to evaluate progress

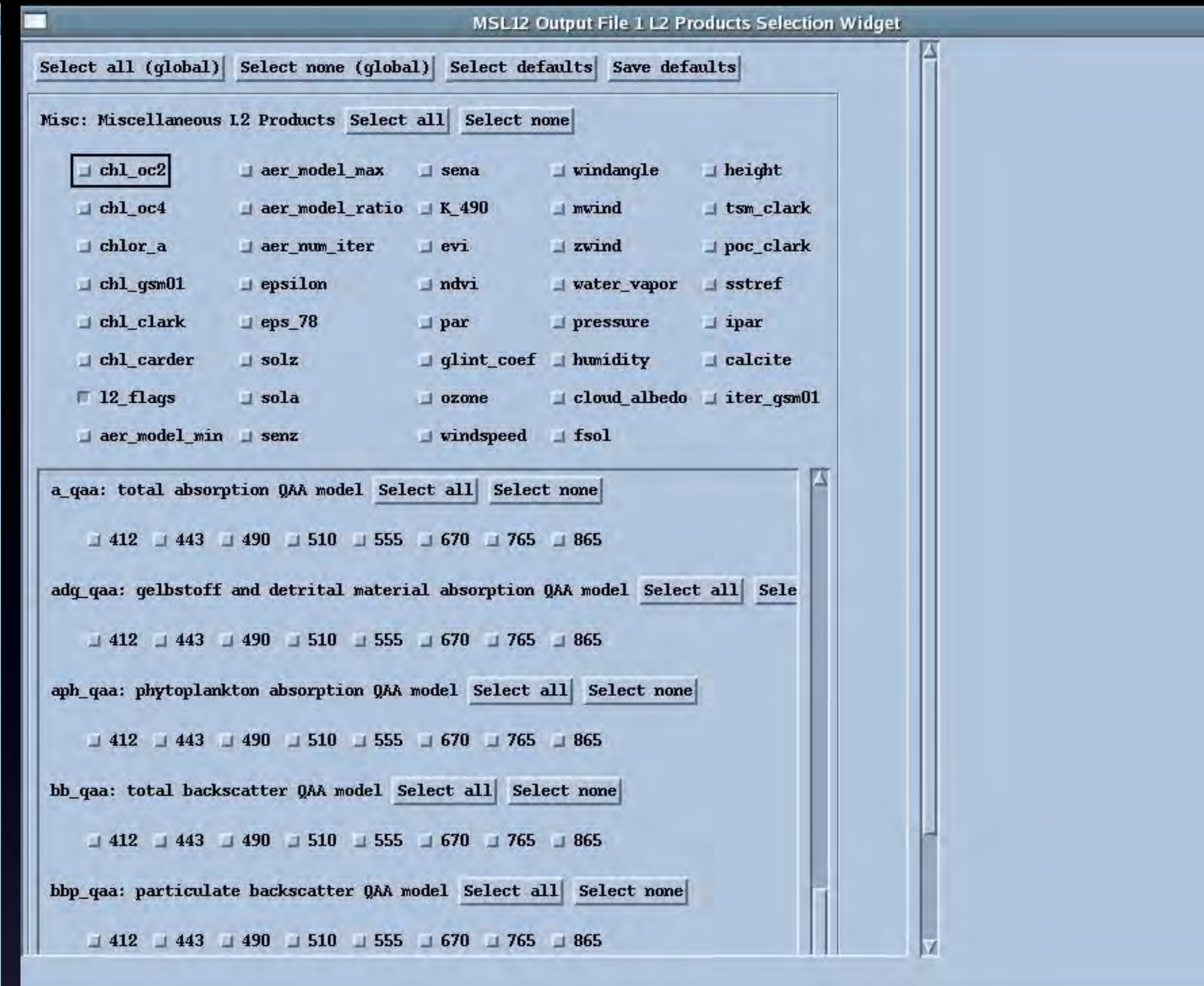
generalized IOP model (GIOP) in l2gen

- specify sensor wavelengths to fit
  - e.g., 412,443,490,510,555
- select a_{ph} form and set params
  - tabulated: λ, $a_{ph}^*(\lambda)$
  - gaussian: λ_0, σ
  - dynamic: Bricaud, Ciotti, Lee
- select a_{dg} form and set params
  - exponential: λ_0, S
  - dynamic: QAA, OBPG
- select b_{bp} form and set params
  - power law: λ_0, η
  - dynamic: HL96, QAA, LS00, Ciotti, Morel
- select rrs[0-] to bb/(a+bb)
  - quadratic
  - f/Q: Morel (tbd: PML, Lee)
- specify inversion method
  - Levenburg-Marquart
  - Amoeba (downhill simplex)
  - Lower-Upper Decomposition
  - Singular-Value Decomposition
- specify output products
  - $a(\lambda)$, $a_{ph}(\lambda)$, $a_{dg}(\lambda)$, $b_b(\lambda)$, $b_{bp}(\lambda)$
    λ = any sensor wavelength(s)
  - C_a (given a_{ph}^* at λ_0)
  - η, S (dynamic model params)
  - internal flags

Other inversion methods

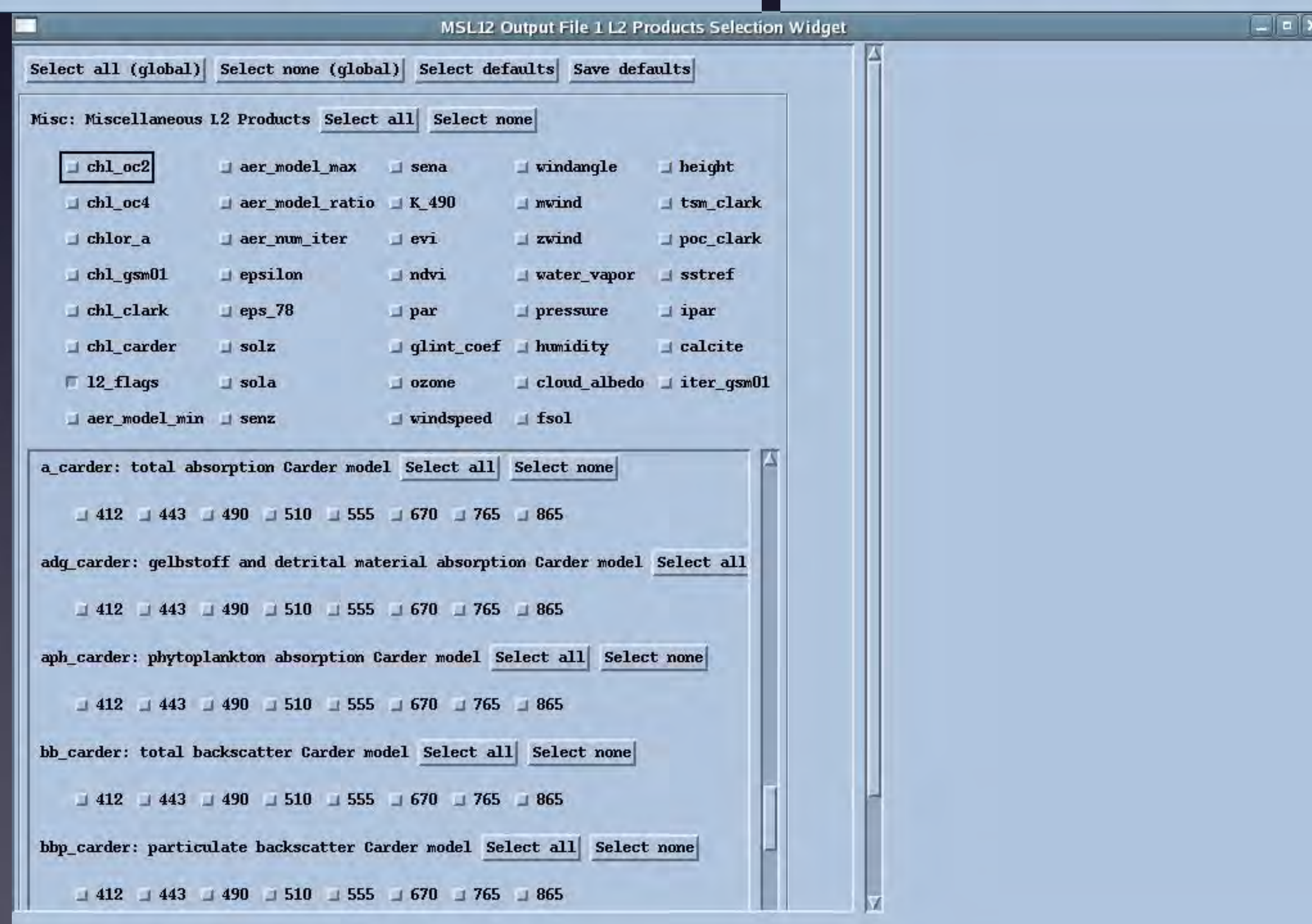
- Genetic algorithms
- Other Neural Network methods (e.g. radial basis algorithm)
- Look Up Tables
- Ambiguity issues





GSM

QAA



Carder

Ocean color products
(REASoN, MEaSUREs and EOS programs)

Product	Link to biogeochemistry	Algorithm
Chlorophyll-a	Phytoplankton biomass; Primary Production	GSM semi-analytical model
$a_{cdm}(\lambda)$	Photochemistry; Heterotrophic production; Light budget	GSM & QAA algorithms
$a_{ph}(\lambda)$	Physiology and type of phytoplankton; Primary Production; trophic state	GSM & QAA algorithms
$b_{bp}(\lambda)$	Particulate material; POC	GSM & QAA Loisel et al. (2006)
S - $a_{cdm}(\lambda)$ spectral slope	Photochemistry, CDOM origin & bleaching history	GSM semi-analytical model QAA algorithm
η - $b_{bp}(\lambda)$ spectral slope	Particle size distribution Export flux	Loisel et al., (2006)
$K_d(\lambda_{UV})$	Light Budget, Photochemistry	Siegel et al. (2007)
Phytoplankton Functional Types	Primary Production Carbon fluxes	Alvain et al. (2004, 2006)
Net Primary Production	Primary Production Carbon fluxes	VGPM & CbPM
Merged products (chl, $a_{cdm}(443)$, $b_{bp}(443)$)	Phytoplankton biomass, Primary & secondary production, Particulates, POC, Photochemistry	Maritorena & Siegel (2005)

3 main families of techniques:

- statistical techniques (1D, 2D):
  - averaging (AVG)
  - blending (BA)
  - objective analysis (OA)
  - EOF-based
  - wavelet analysis (WA)
  - machine learning (MLA)

(Gregg & Conkright, 2001, Kwiatkowska & Fargion, 2003, Pottier et al., 2006, 2008, Saulquin et al., 2011)

- optically-based techniques (OB):
  - uses full spectral information

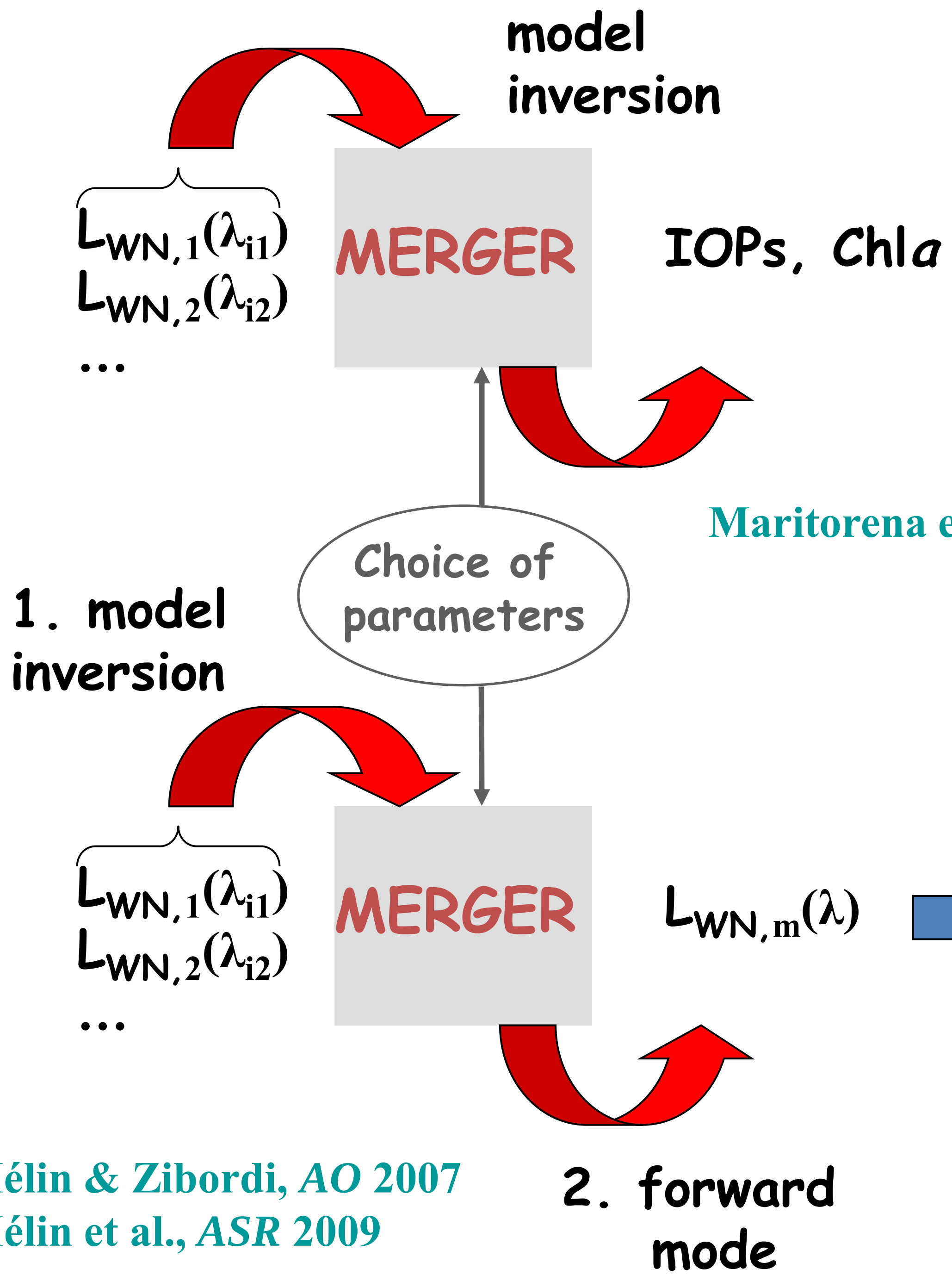
(Maritorena et al., 2005, 2010, Mélin & Zibordi, 2007, Mélin et al., 2009, 2011)

- numerical model-based methods
  - assimilation in BGC models

(Gregg 2008)

…. Different approaches can/must be combined

WP2800: OB Techniques



Provides only merged IOPs + Chla

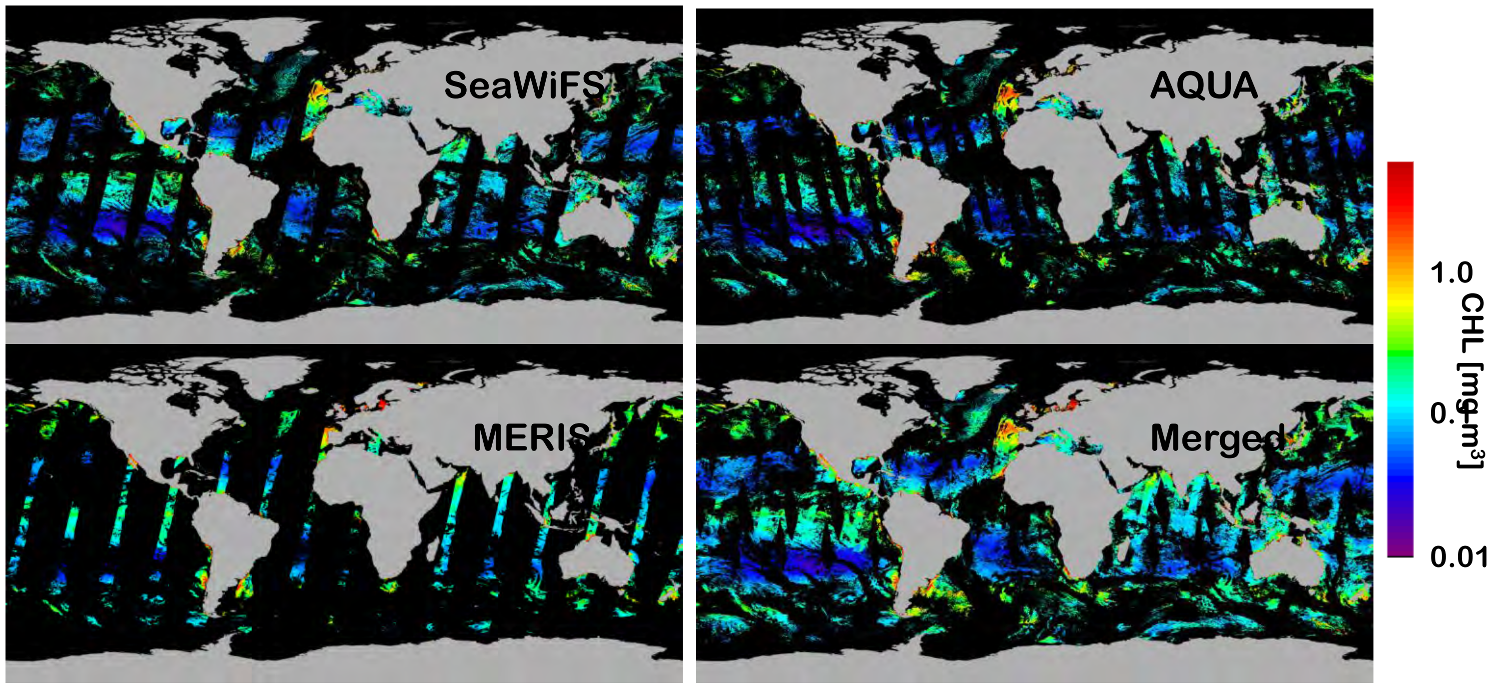
Maritorena et al., *RSE* 2005, 2010

Provides merged R_{RS},
but with implicit IOPs
except with multiple solutions

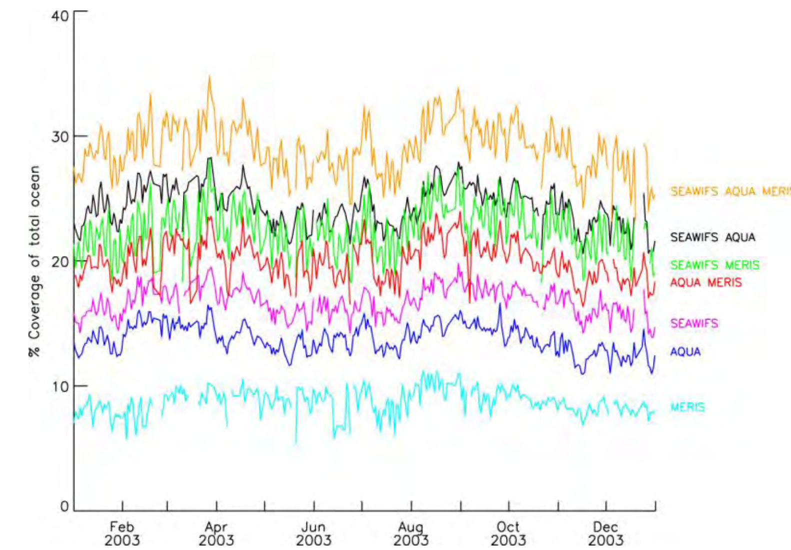
Mélin & Zibordi, *AO* 2007
Mélin et al., *ASR* 2009

'Choice' of bio-optical model imbedded in R_{RS}

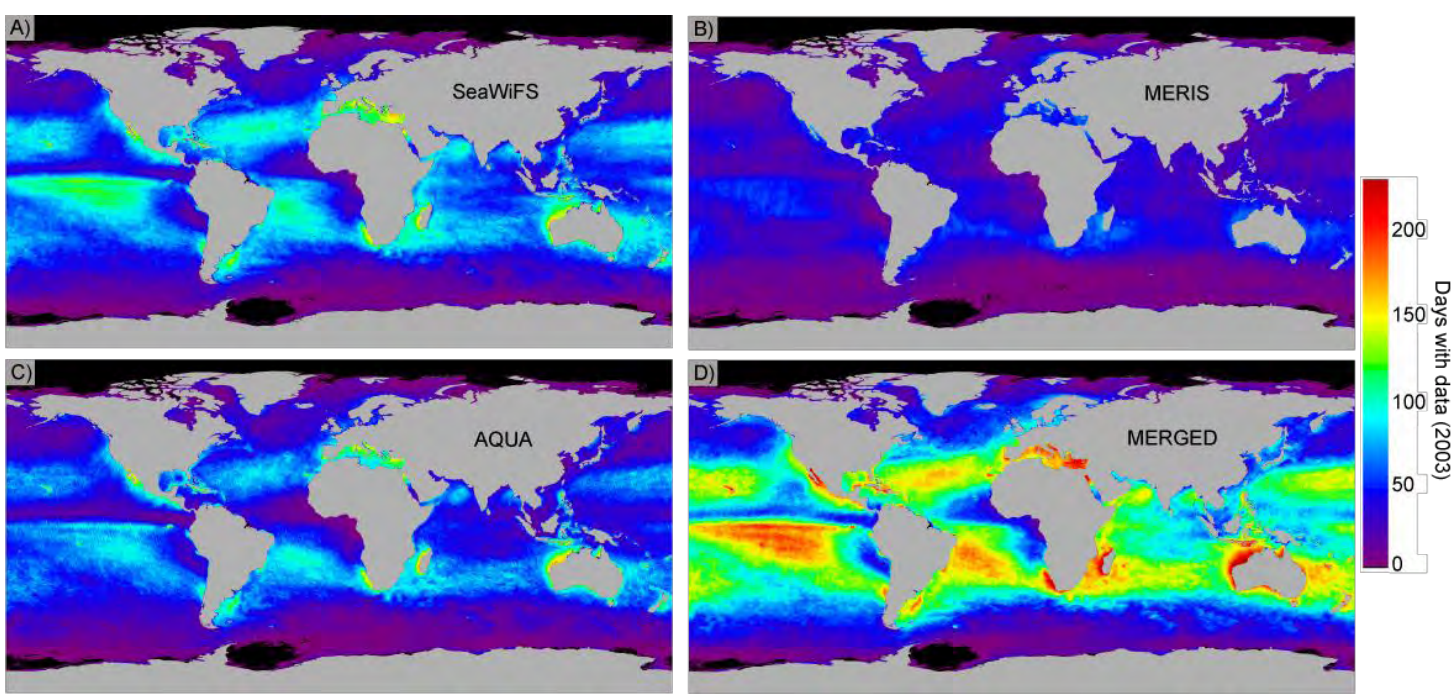
Merged data sets



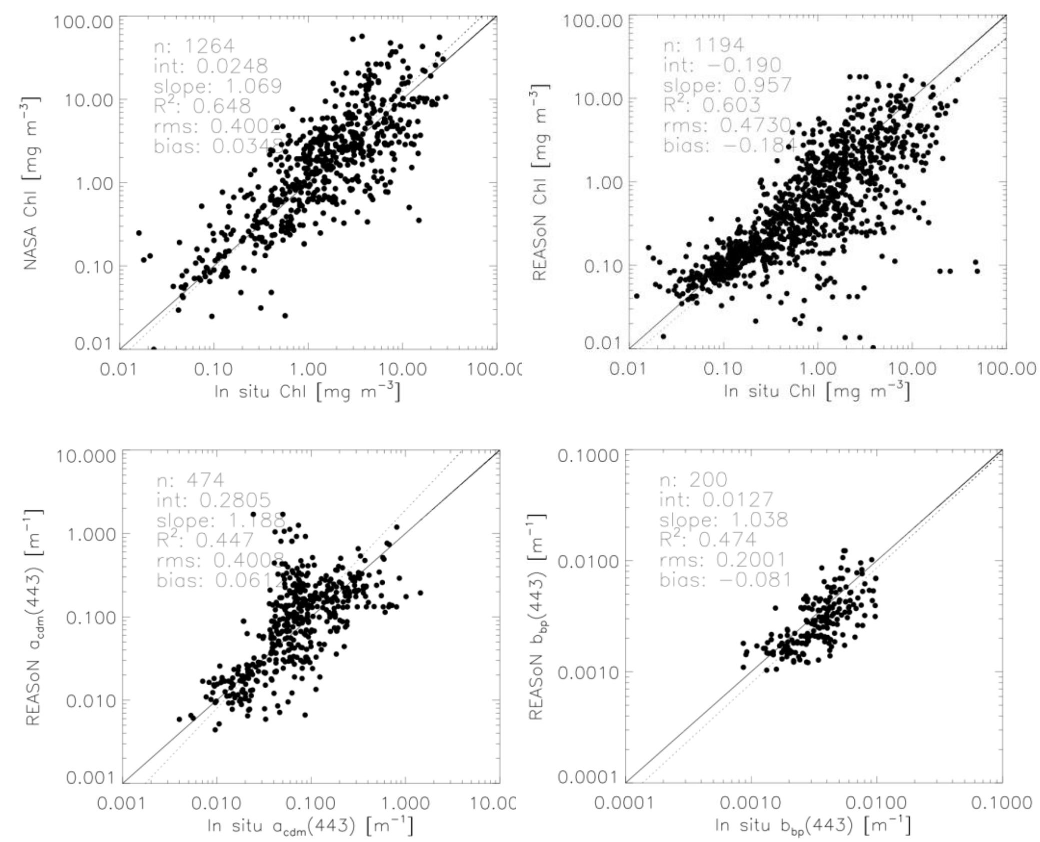
Spatial coverage



Temporal coverage



Matchups



ftp:ftp.oceancolor.ucsb.edu/pub/org/oceancolor/REASoN/
OPeNDAP server: http://dap.oceancolor.ucsb.edu/cgi-bin/nph-dods/data/oceancolor/

Section 4

Algorithm implementation issues, regional and class based algorithms

Global vs Regional algorithms

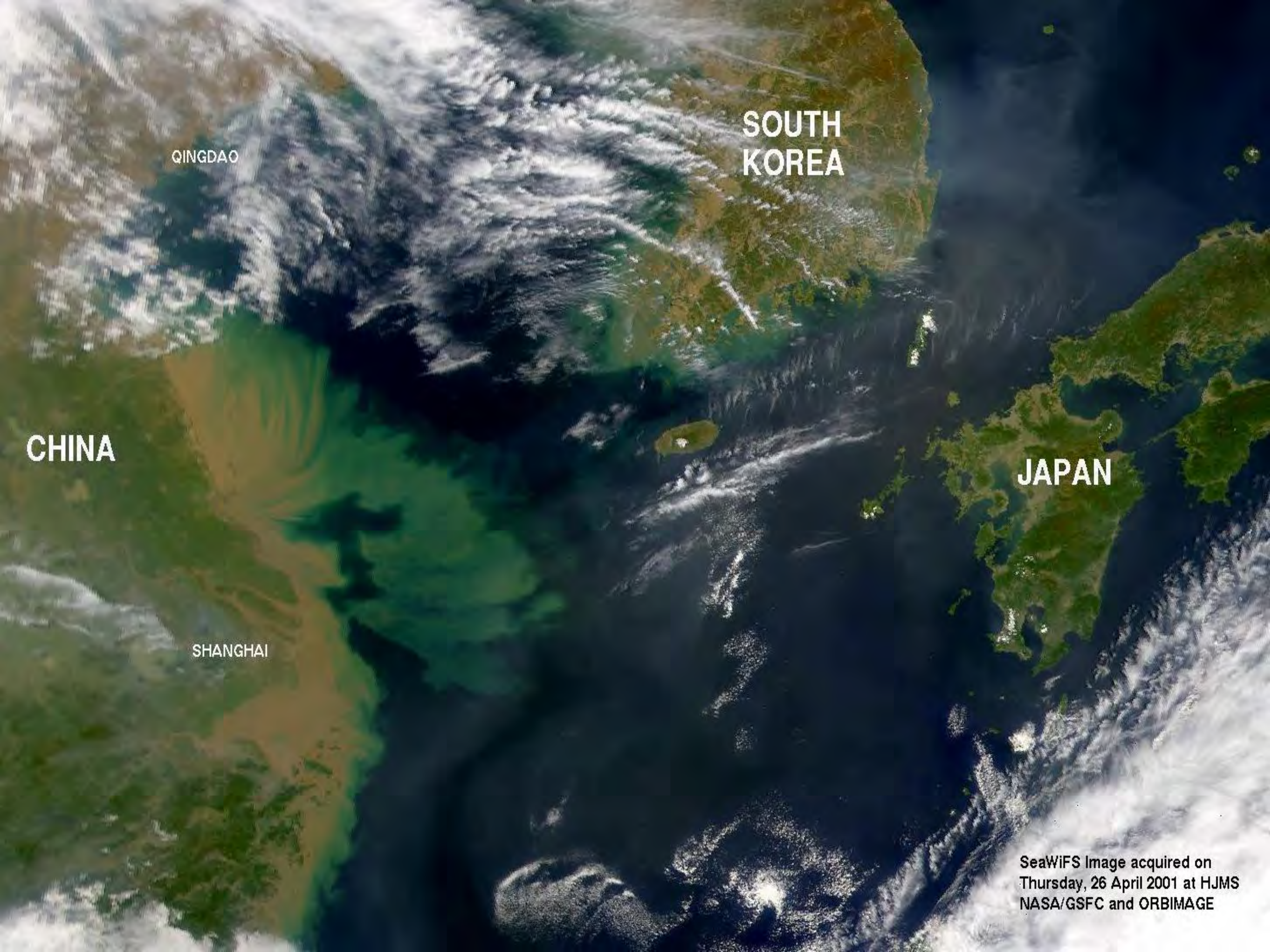
- Global dataset now consist of multiple “products” geophysical variables and IOPs
- They have the advantage of being routinely produced by the space agencies globally
- They also generally have “mean” uncertainties associated with them
- It may be that one of the available products perform well in your region
- Alternative is to produce your own regional algorithm, with local parameterisation

Data Requirements for Regional Algorithms (from “cheap” to expensive”)

1. Simple validation of standard SA products: Fluorometer -> HPLC for Chl a
2. Spectrophotometric estimate of aph and aCDOM some estimate of bbp
3. Validation of nLw, Rrs (15-25 K$)
4. Full in-water IOP AOP dataset (80K$ equipment)
5. Full CAL/VAL IOP AOP programme with fully traceable persistently calibrated instruments > 150K$ /year

Regional vs. Class-Based

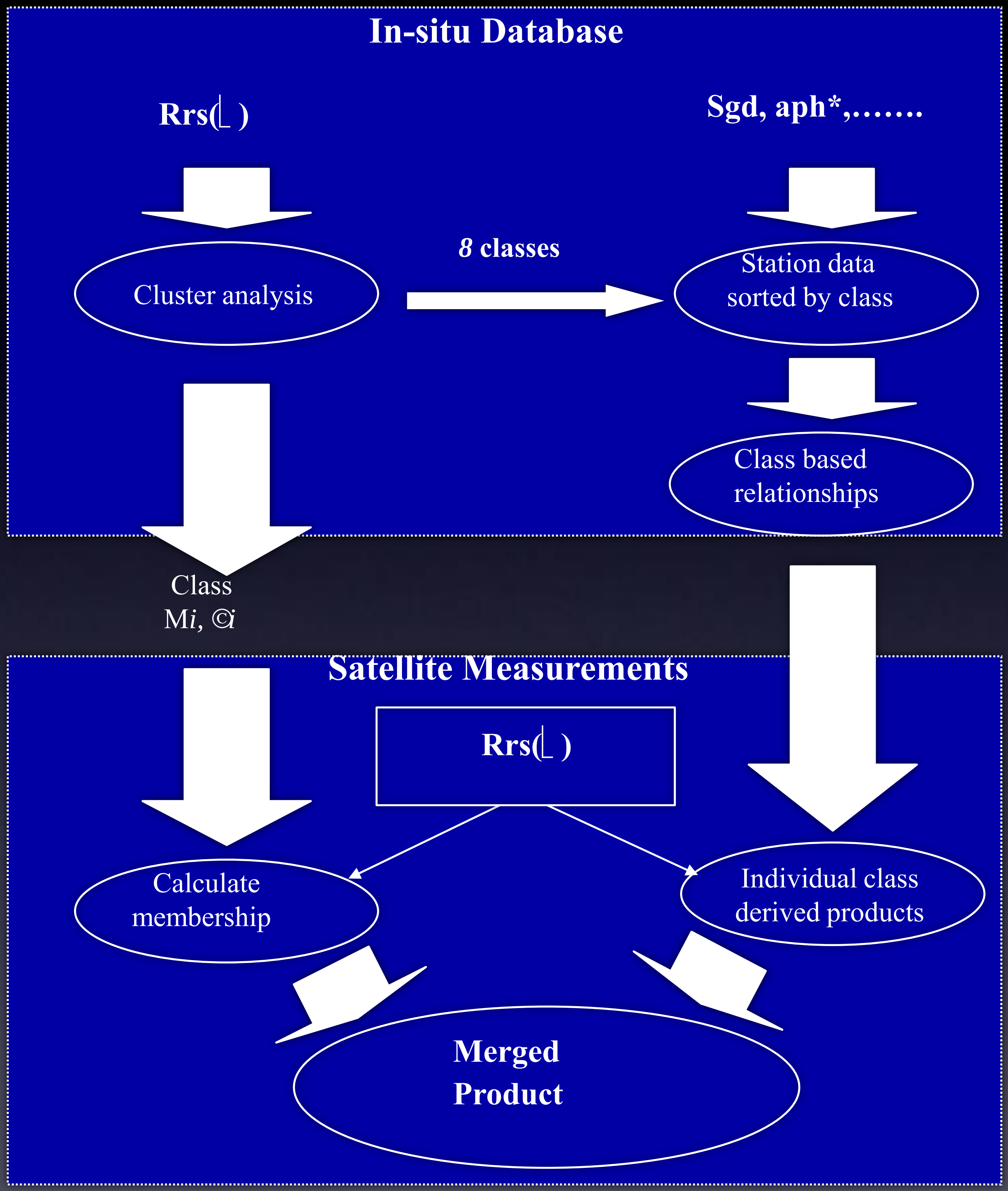
	Regional	Class-based
Advantages	1. Explicitly linked to locally measured in-situ data 2. May be “simpler” 3. Accounts for physiological differences	1. Generic, “global”, can be generalised 2. Can be used as a tool to identify “black holes” 3. Seamless transitions 4. Continuous improvements through additional on in-situ data
Disadvantages	1. Explicitly link to locally measured in-situ data – not generalized 2. May result in regional discontinuities	1. More complicated to implement 2. Computational more expensive – not much!


QINGDAO
SOUTH KOREA
CHINA
JAPAN
SHANGHAI
SeaWiFS Image acquired on
Thursday, 26 April 2001 at HJMS
NASA/GSFC and ORBIMAGE

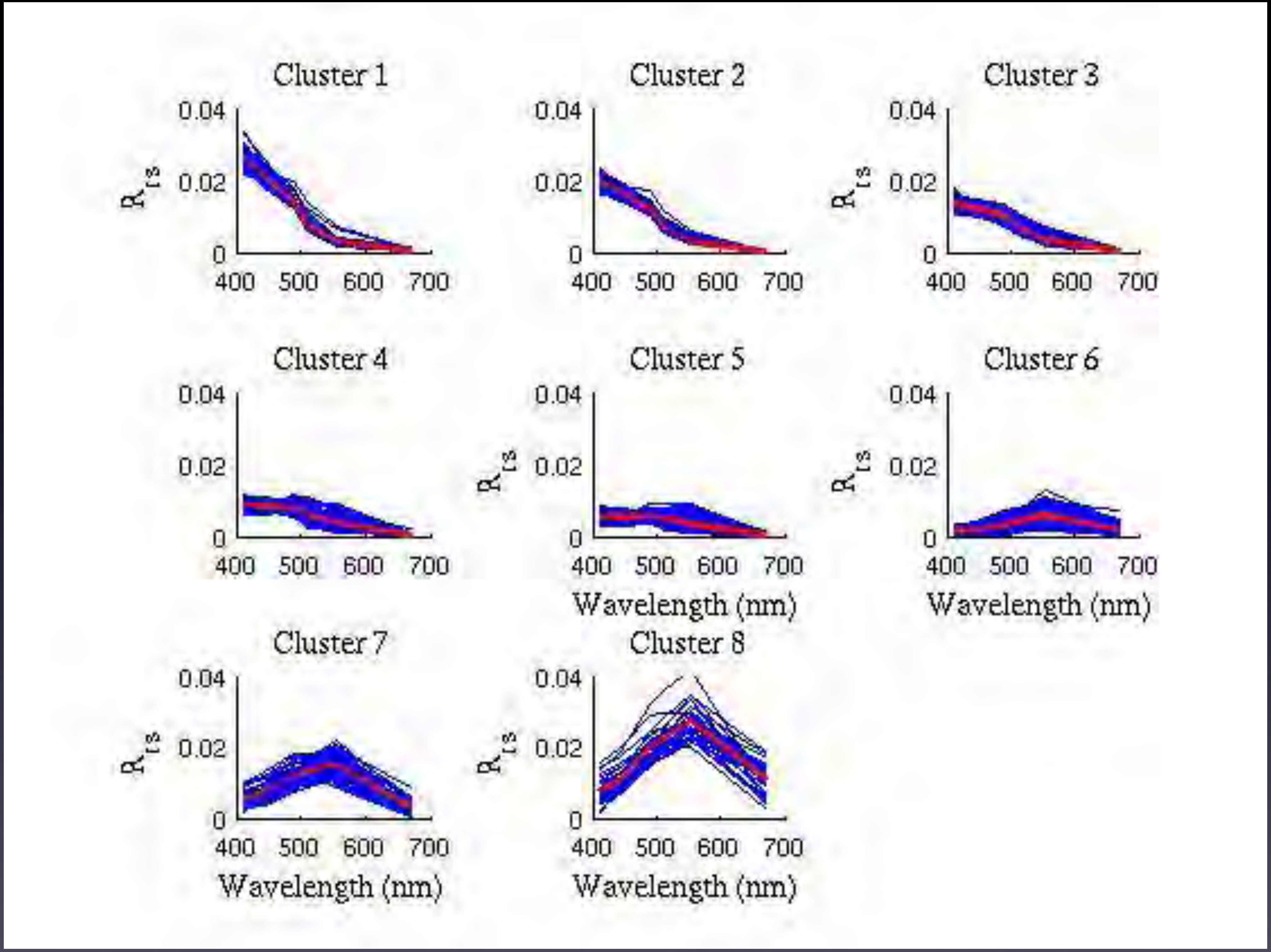
Rationale

- There is necessity to describe a considerable amount of variability in Inherent Optical Property (IOP) subcomponent models.
- This is particularly true, if inversion algorithms are to be applicable at global scale yet remain quantitatively accurate in coastal & shelf seas.
- This is unlikely to be achieved in the foreseeable future, with a single representation of IOP subcomponents.
  - BEAM – Case2R
- The proposed approach is an algorithm framework more than a specific algorithm.

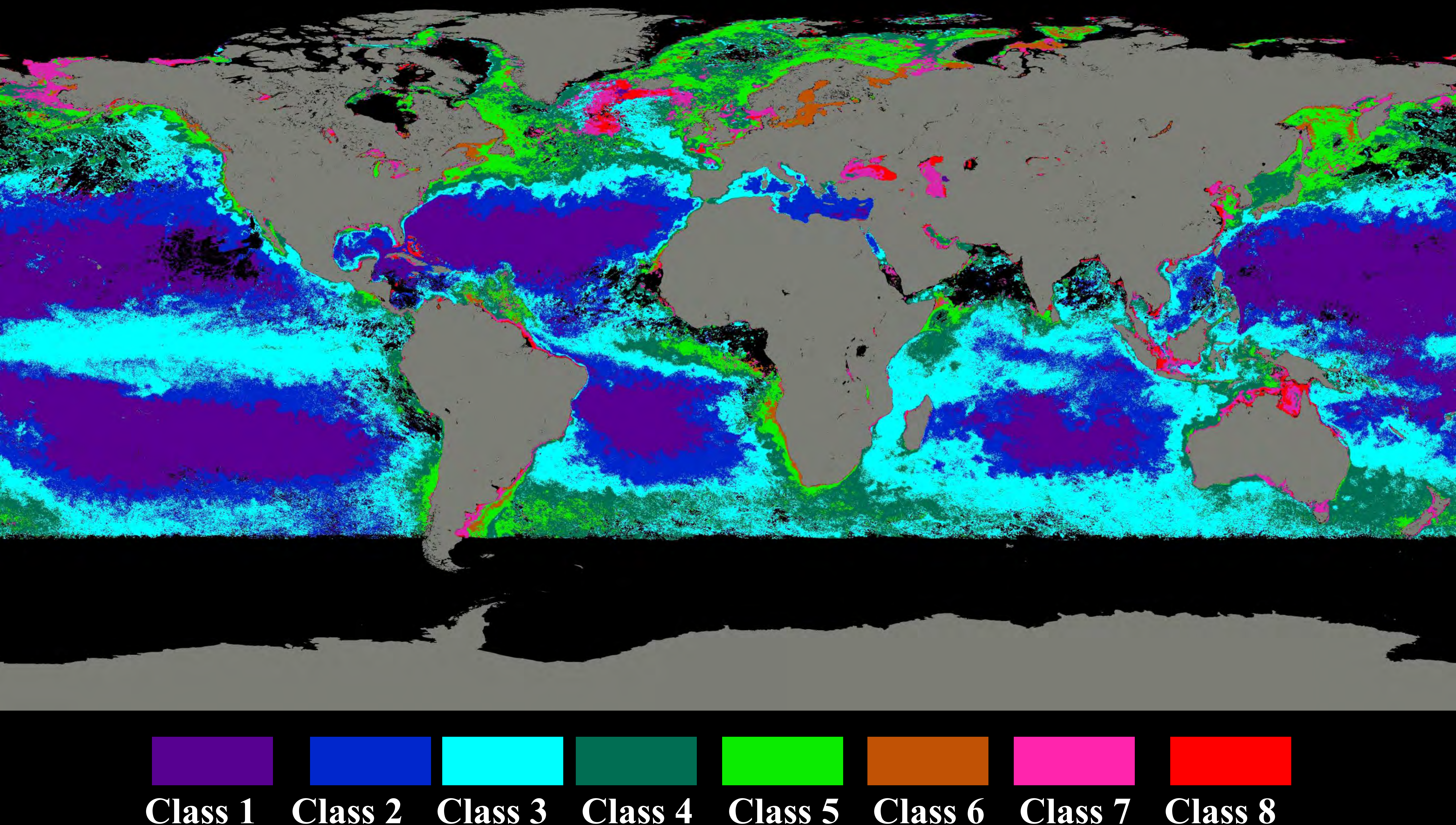
Conceptual Framework for Case based algorithms

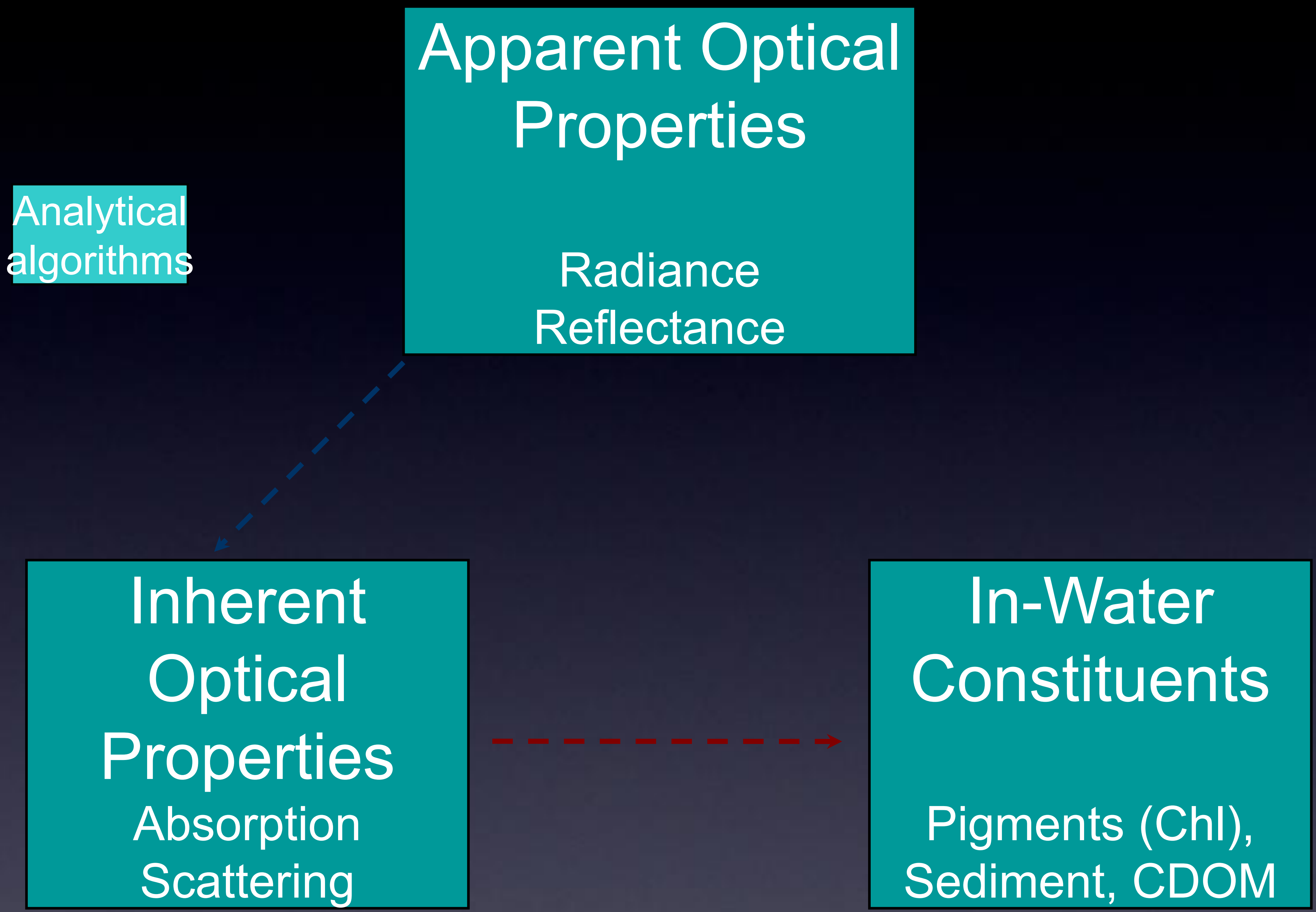


8 objectively identified classes in radiance space



May 2004 SeaWiFS Composite



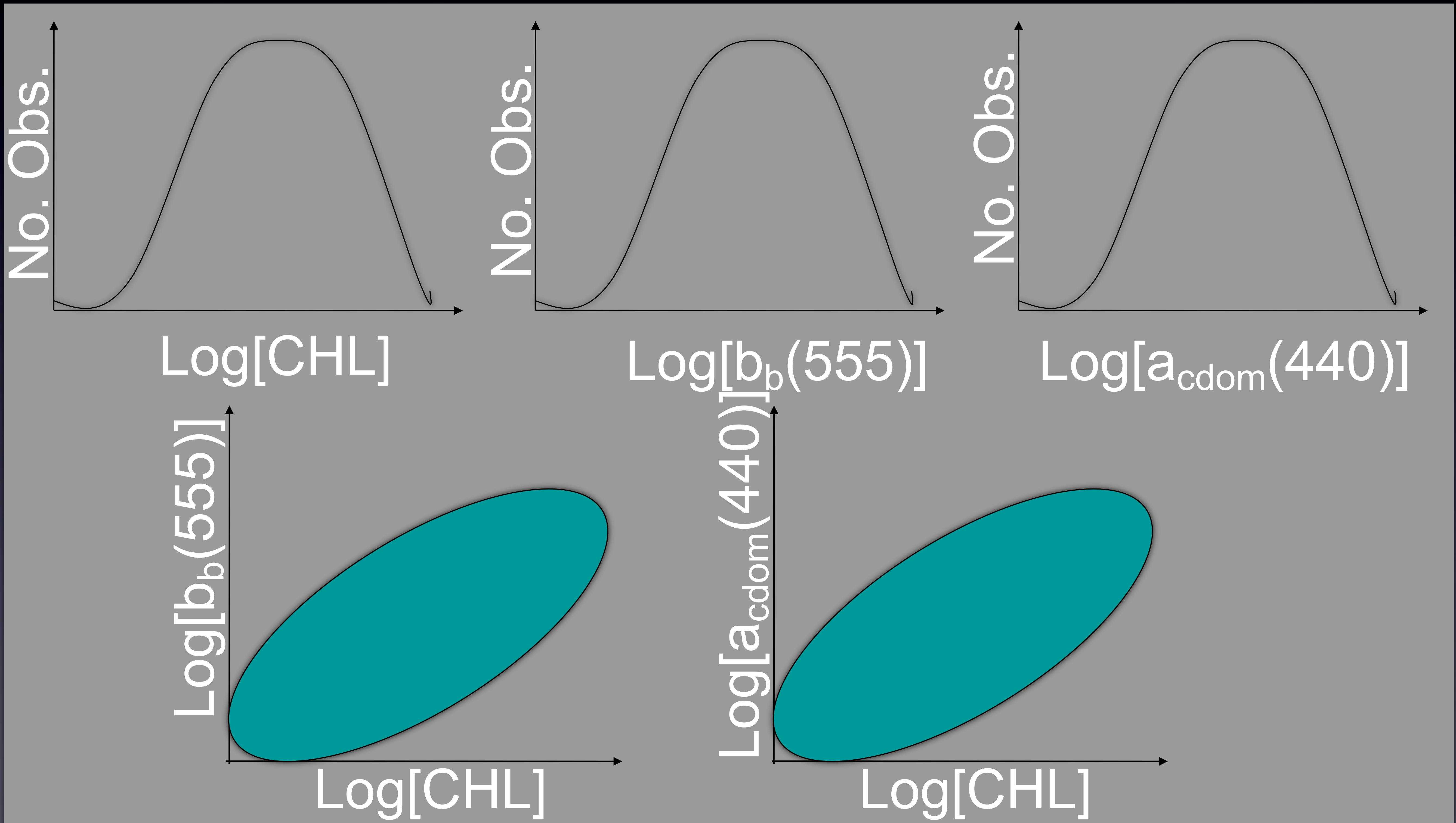


Local parameterization for coastal & inland waters

What to parameterize?

- Variance and Co-variance of Optically Active Constituents
- Parameterising IOP subcomponent models (or fit coefficients – for empirical algorithms)
- Different OWT different inversions method
- Avenue to spatial uncertainty estimates
- Regional value-added products

Distribution functions for in-water constituents



$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda, Chl) + a_d(\lambda, TSS) + a_{cdom}(\lambda, CDOM)$$

$$a(\lambda) = a_w(\lambda) + Ac(\lambda)[Chl]^{Bc(\lambda)} + [a_{cdom}(440)]\, e^{-S(\lambda-440)}$$

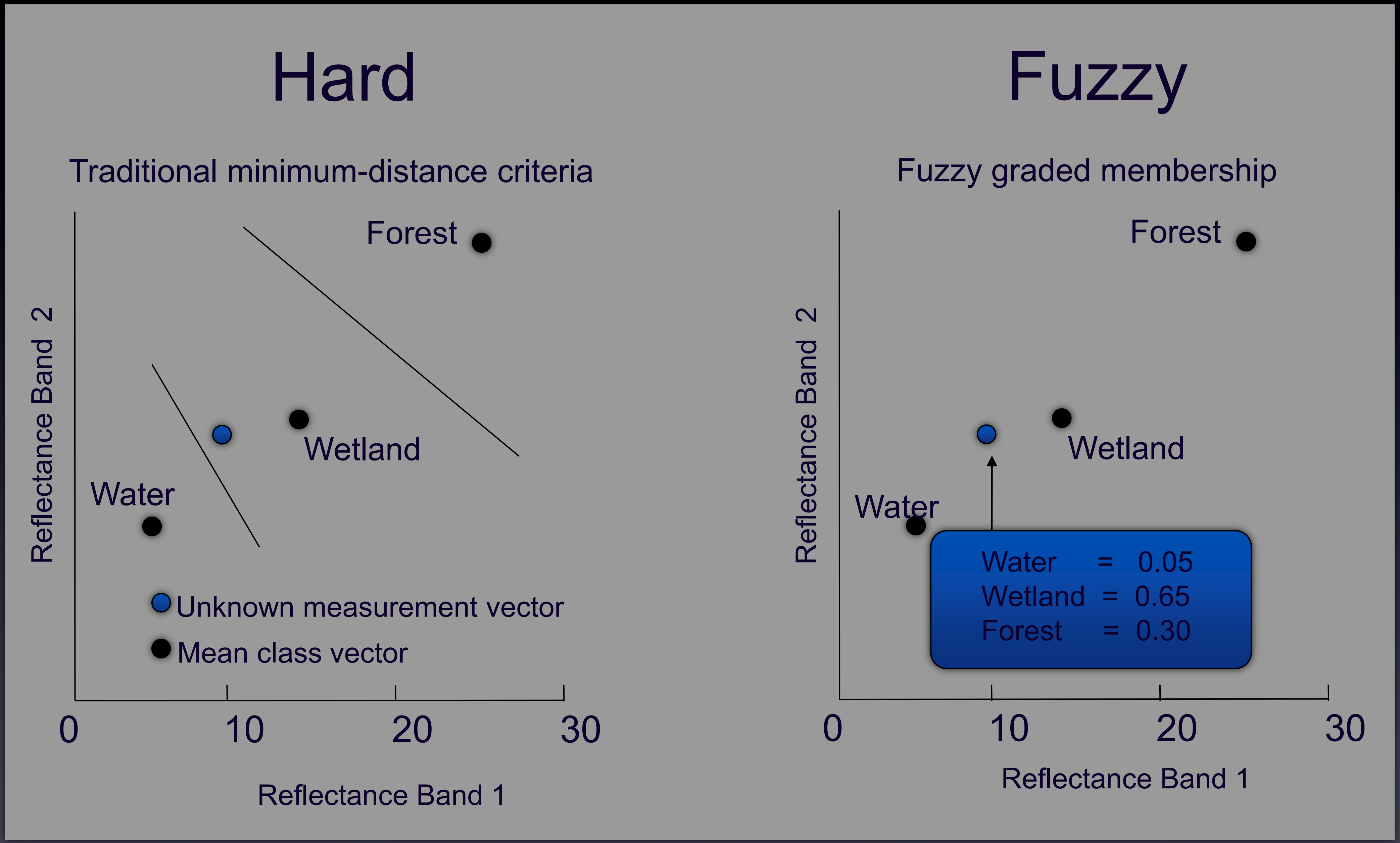
$$\mathrm{b_b}(\lambda) = \mathrm{b_{bw}}(\lambda) + \mathrm{b_{bp}}(\lambda, \mathrm{Chl}, \mathrm{TSS})$$

$$b_b(\lambda) = b_{bw}(\lambda) + [b_{bp}(555)]\, [555/\lambda]^{Y}$$

Methods for Class-based algorithms

- Novelty Detection (D’Alimonte 2002)
- Fuzzy Logic (Moore et. al. 2001)
- Lubac et. al. 2007 - EOF & PCA

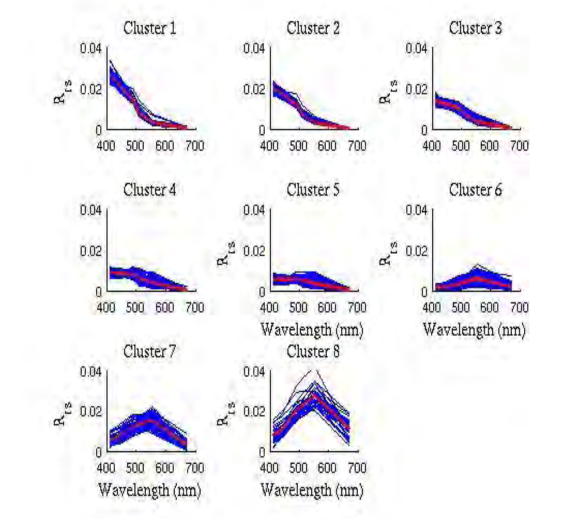
Fuzzy versus Hard classification

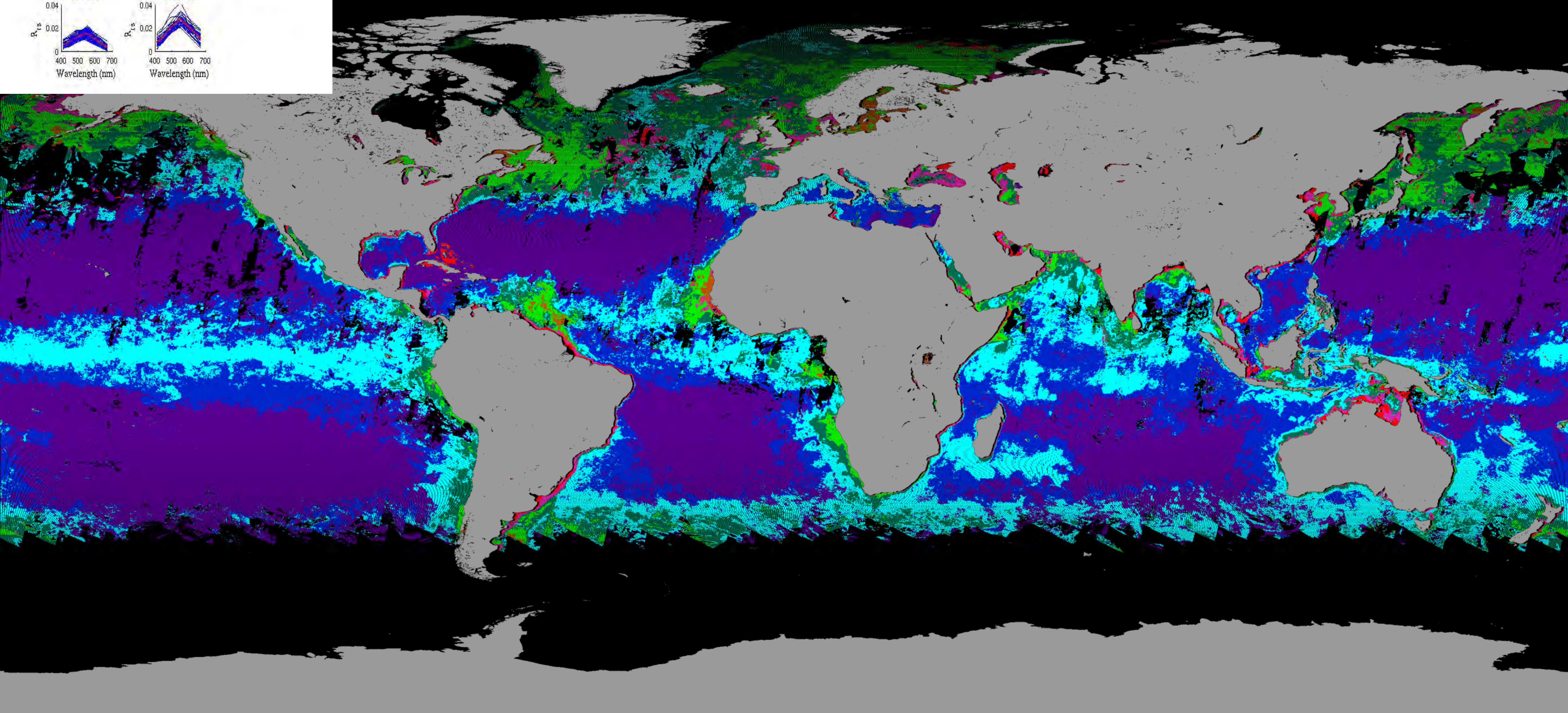


Advantages of fuzzy logic defined provinces

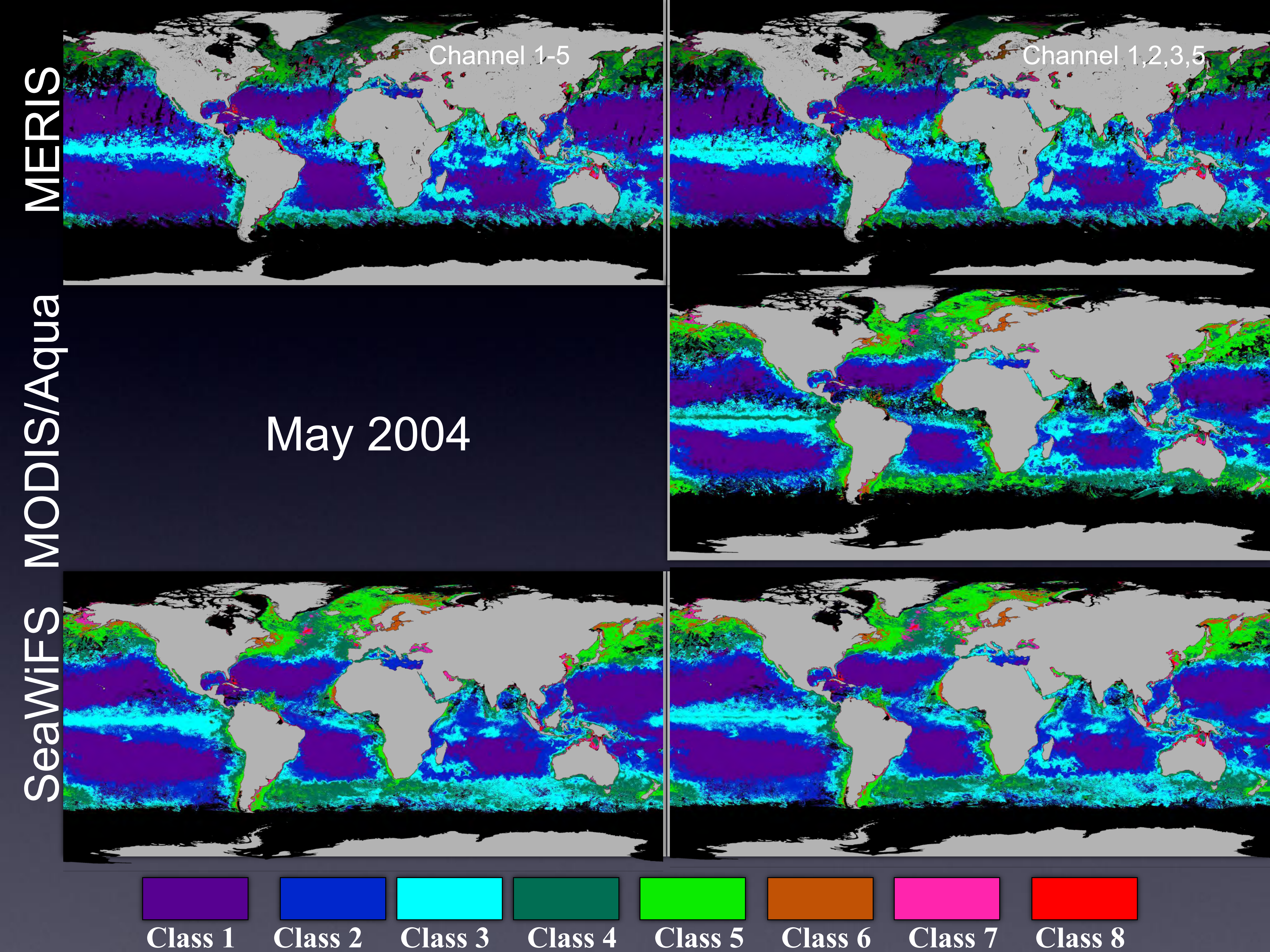
- They allow for dynamics both seasonal and inter-annual in the optical properties of a given region.
- They address the issue of transitions at the boundaries of provinces (through the fuzzy membership function of each class) thus resulting finally in the seamless reconstruction of a single geophysical product.

May 2004 MERIS Global Composite

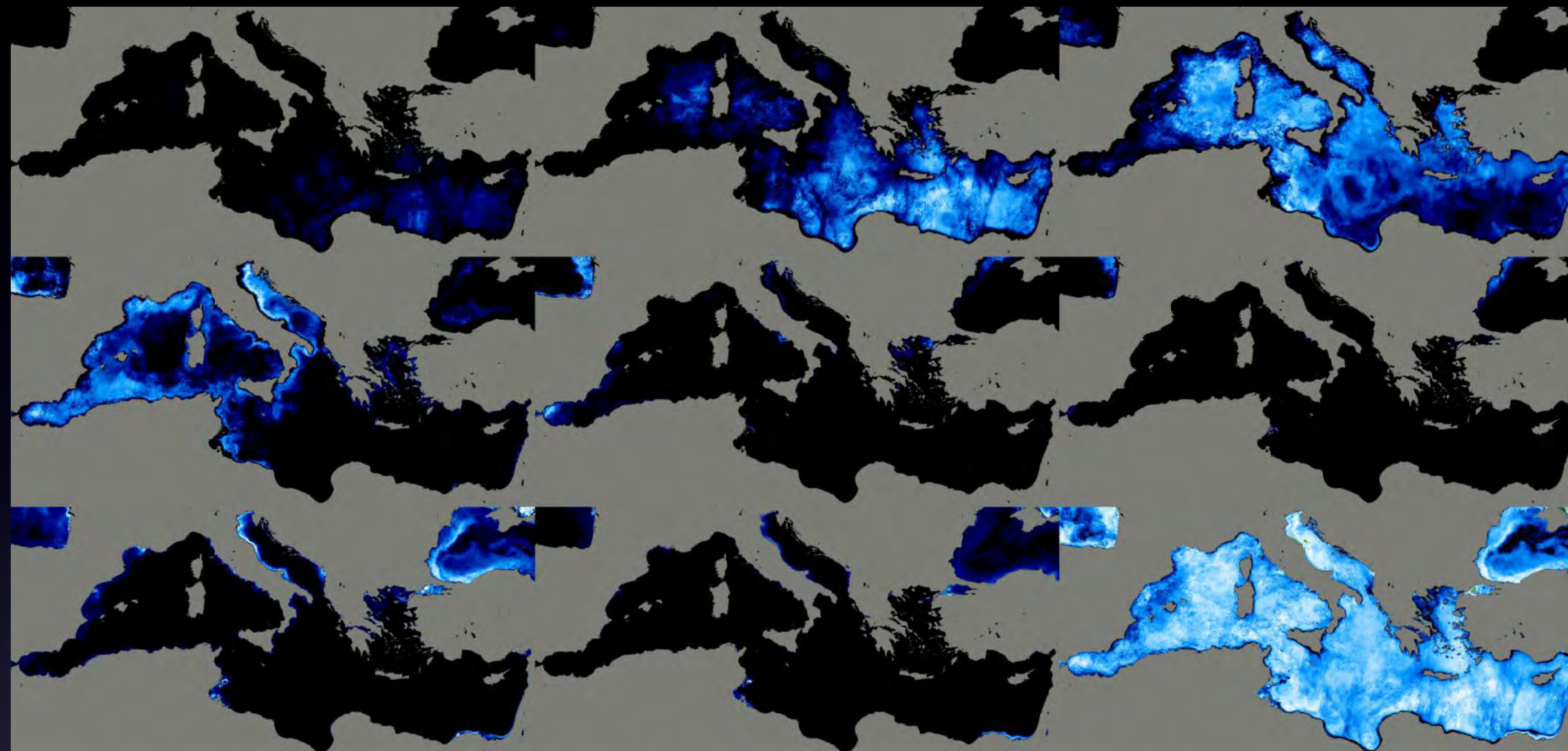


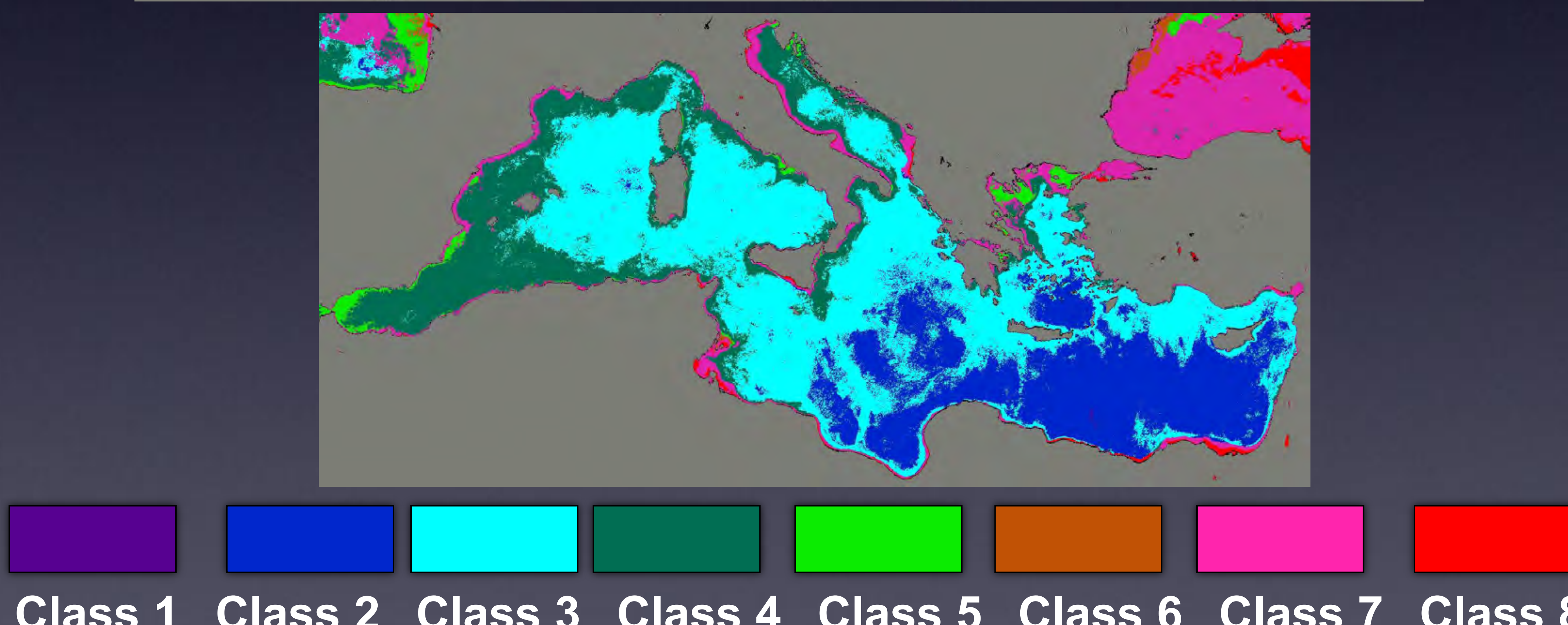


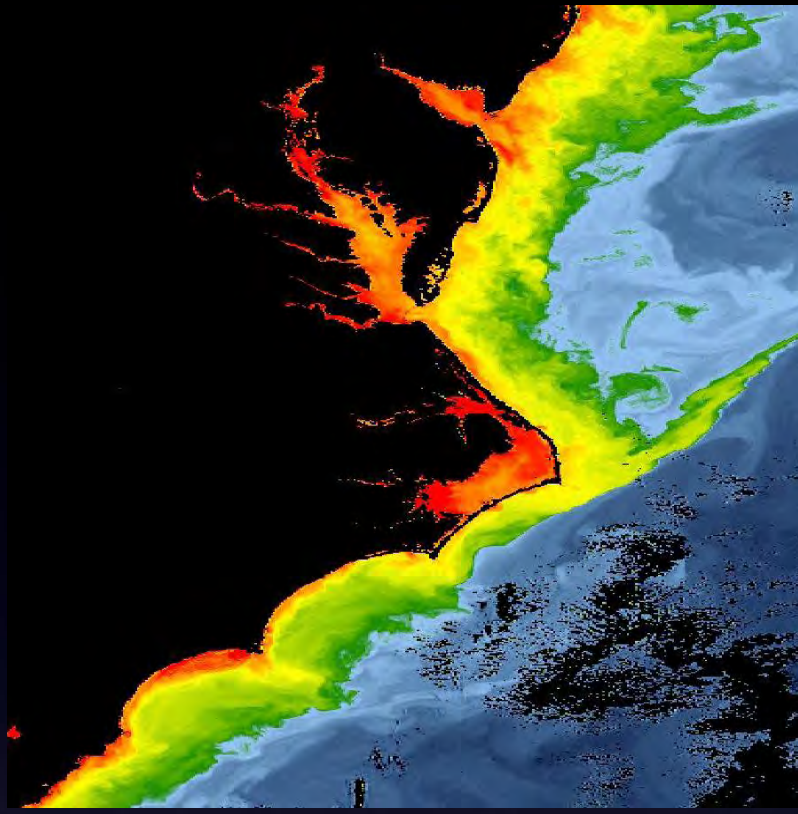
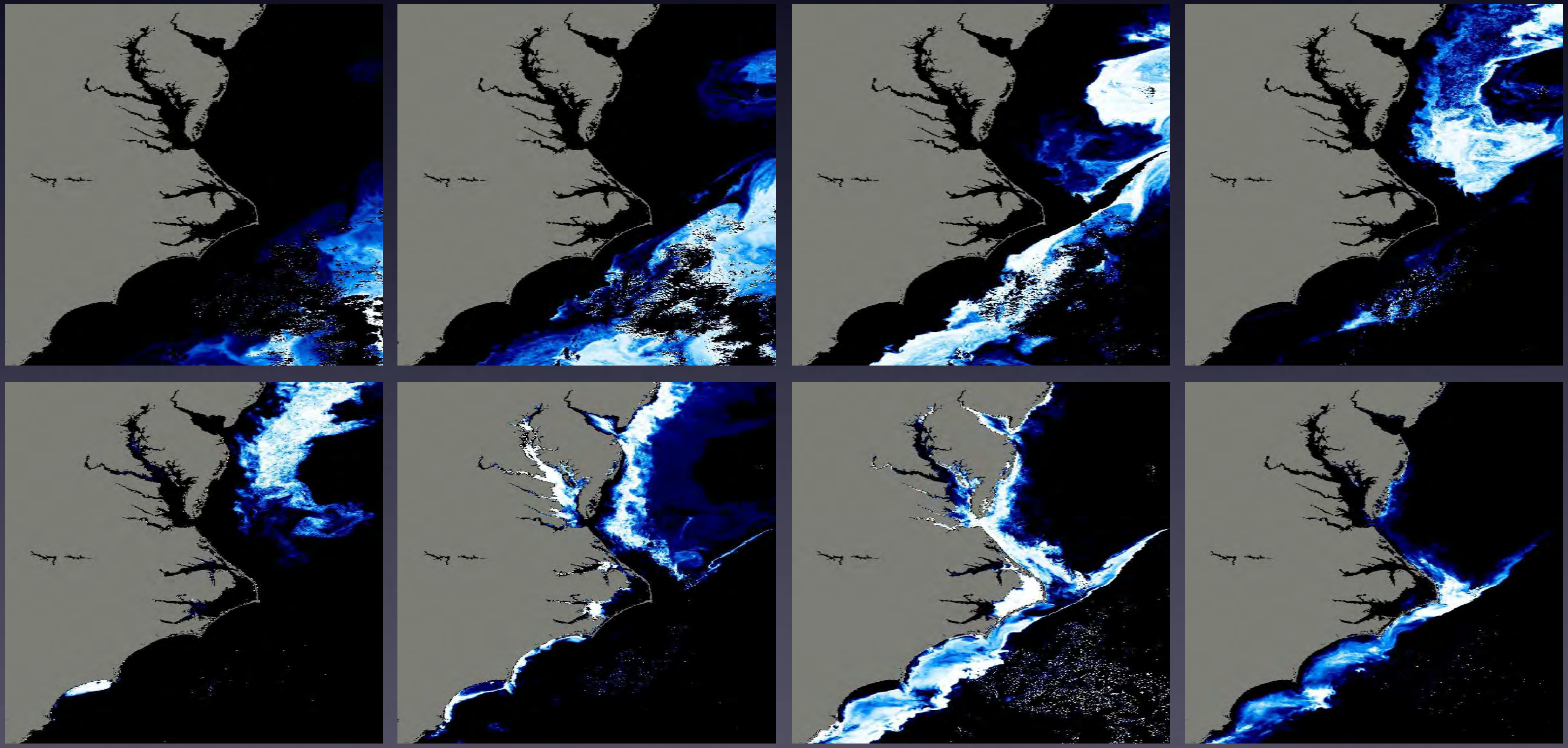
Class 1 Class 2 Class 3 Class 4 Class 5 Class 6 Class 7 Class 8


MERIS
MODIS/Aqua
SeaWiFS
Channel 1-5
Channel 1,2,3,5
May 2004
Class 1
Class 2
Class 3
Class 4
Class 5
Class 6
Class 7
Class 8

High resolution provinces for European Seas
Med May 2004

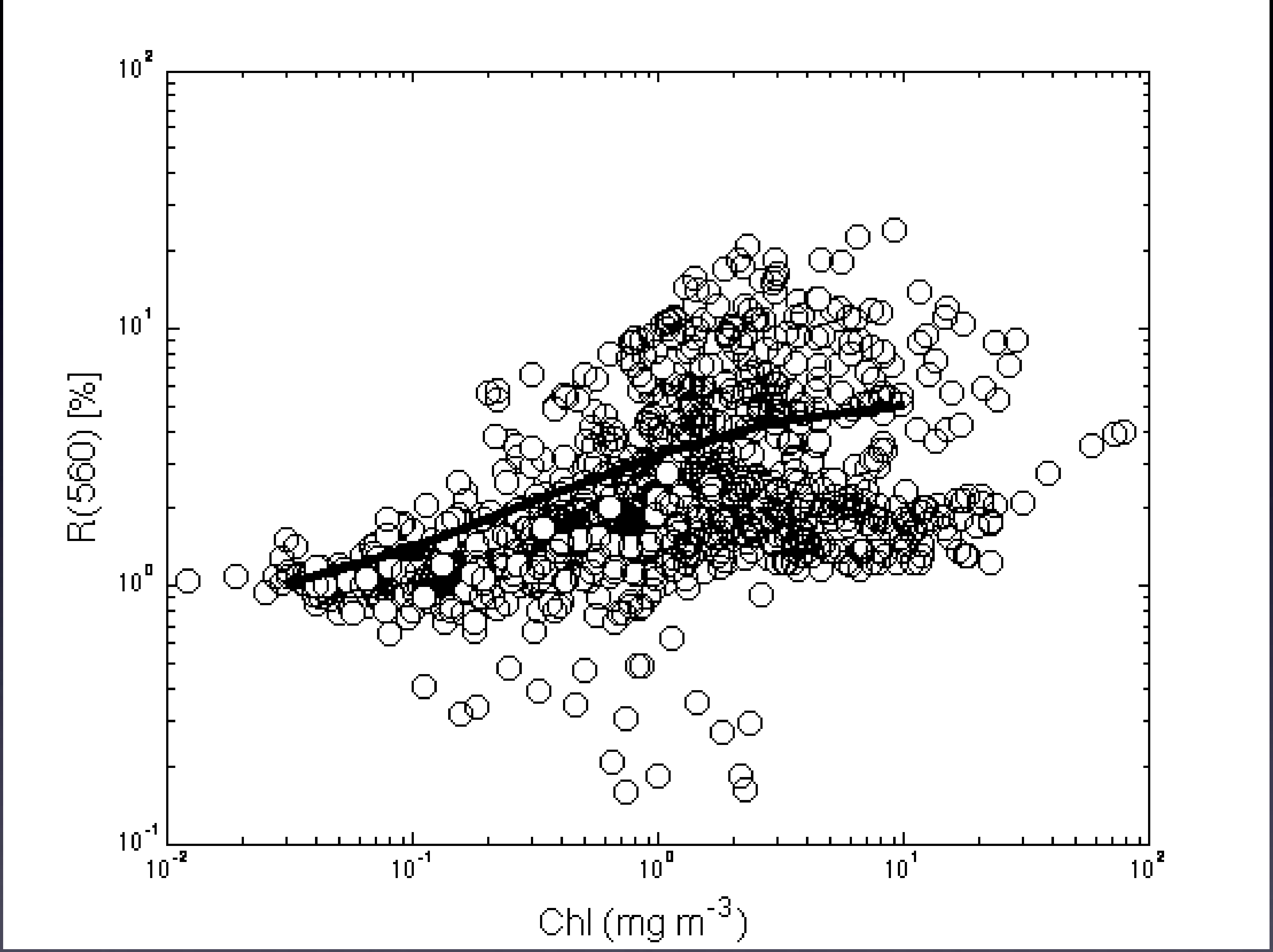





SeaWiFS 23 Sept 1999



Relation to current understanding

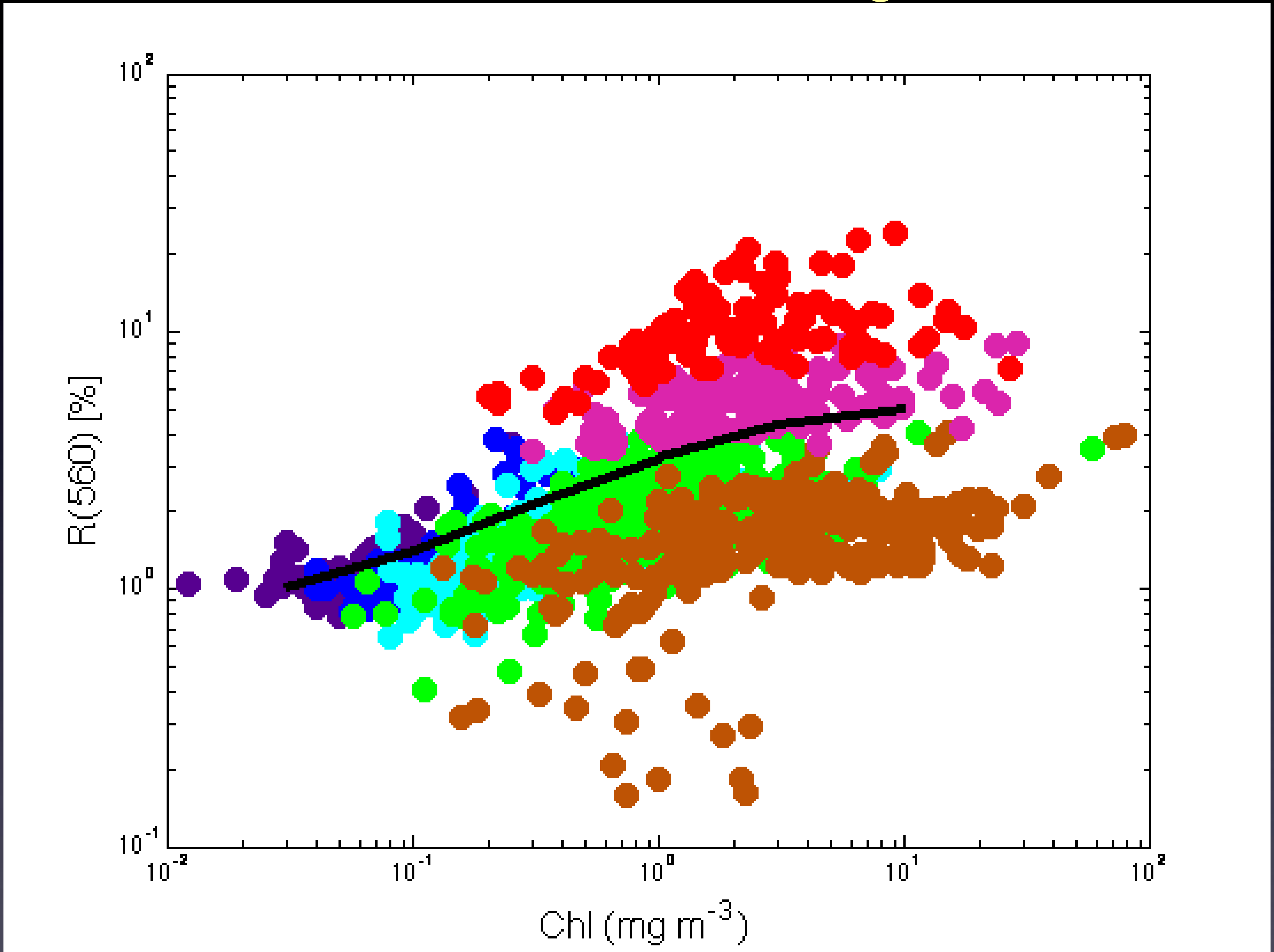
turbid water flag



After Morel and Bélanger 2006

Relation to current understanding

turbid water flag

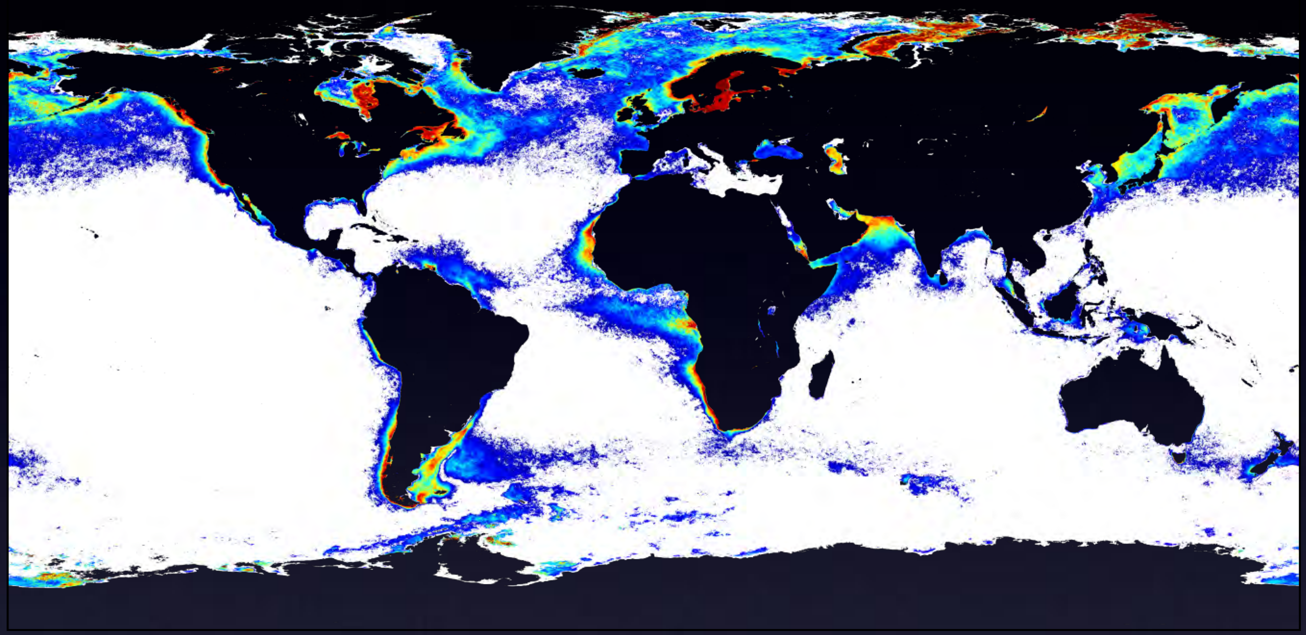


Class 1 **Class 2** **Class 3** **Class 4** **Class 5** **Class 6** **Class 7** **Class 8**

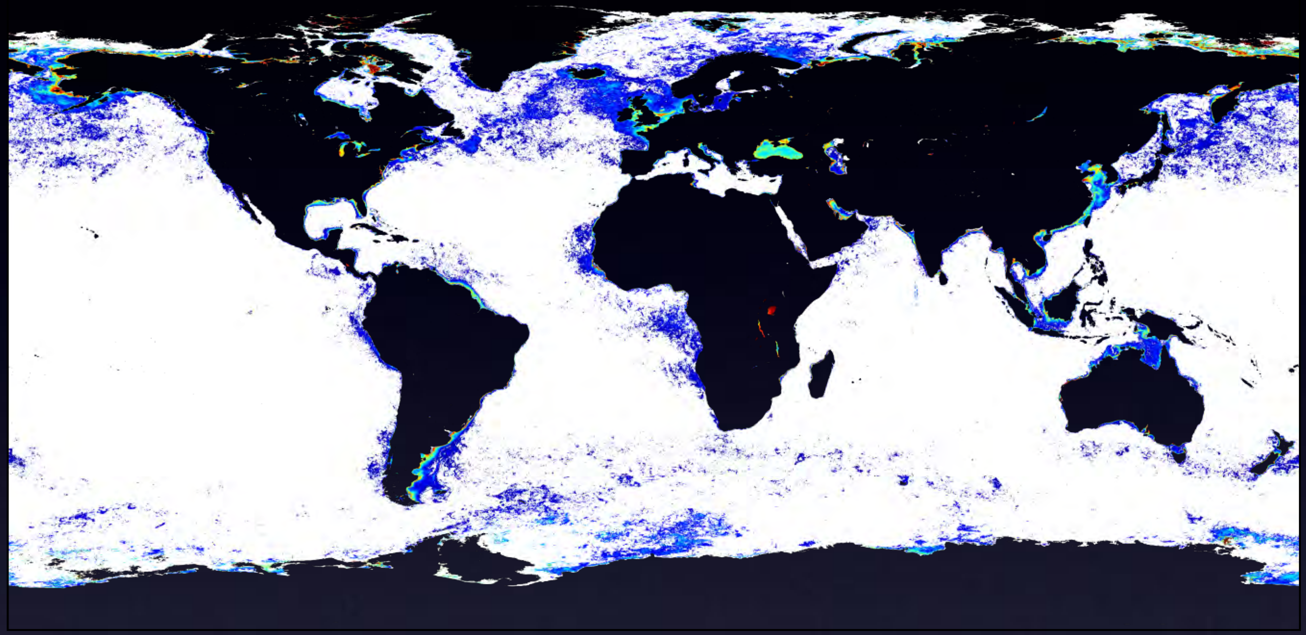
Class persistence

36 month Time-series

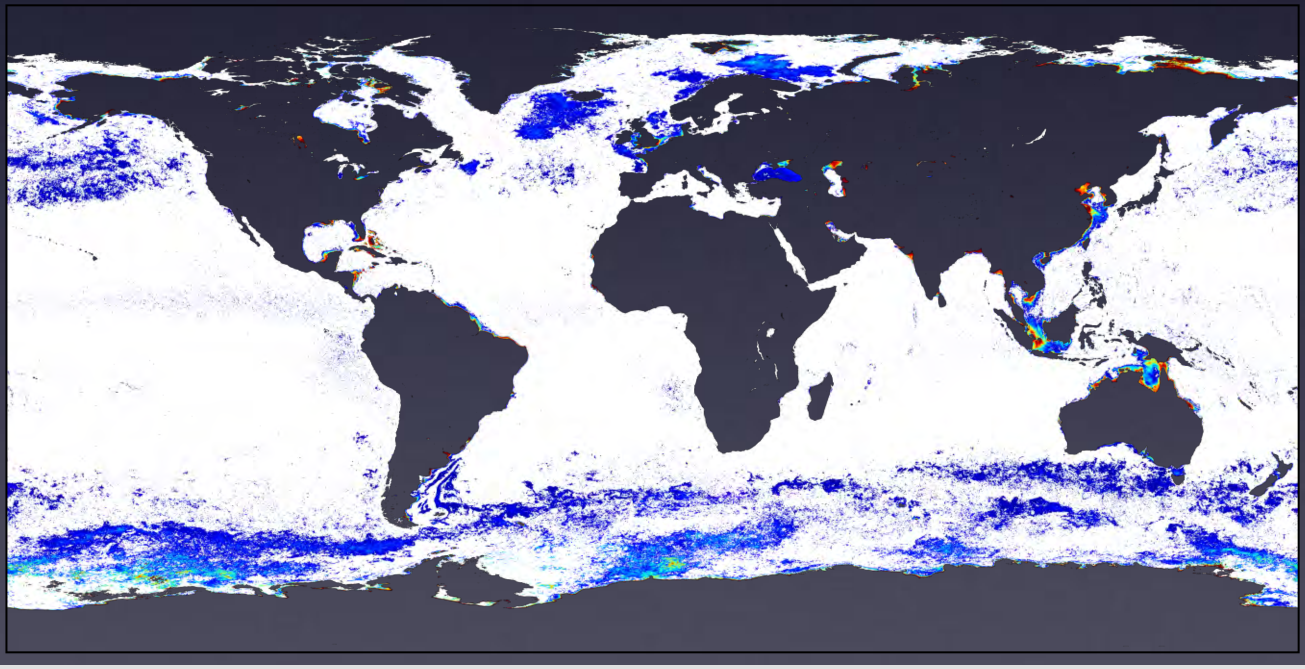
Class 5



Class 6

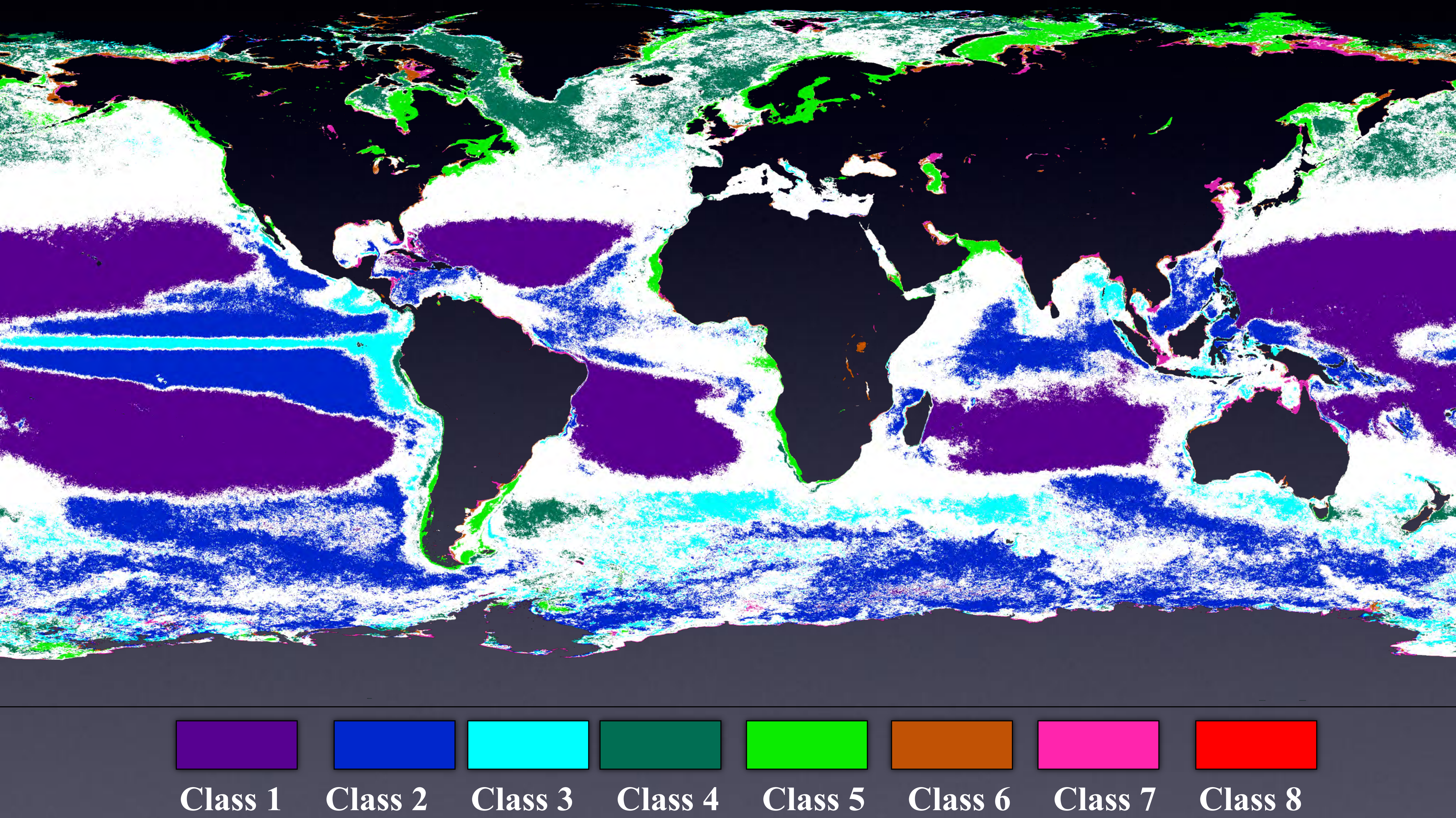


Class 7 & 8



Class Persistence

distribution of classes dominant for more than 70% of observations



Producing the Uncertainty Map

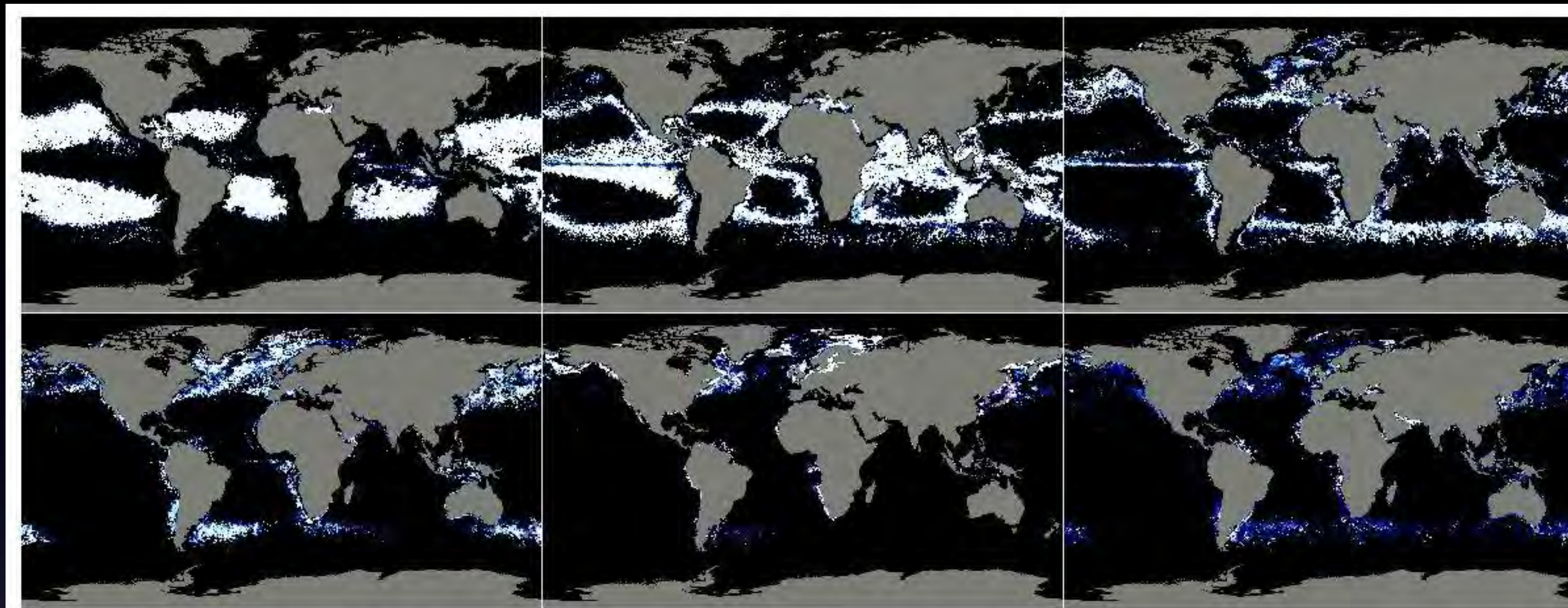
For each pixel,

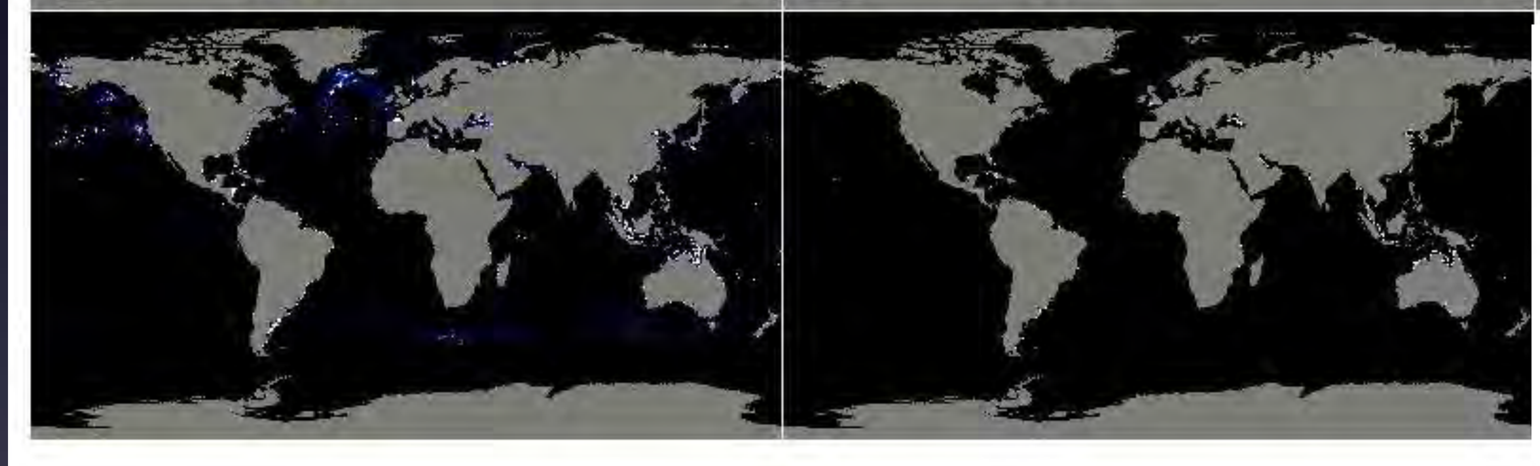
©f_i *

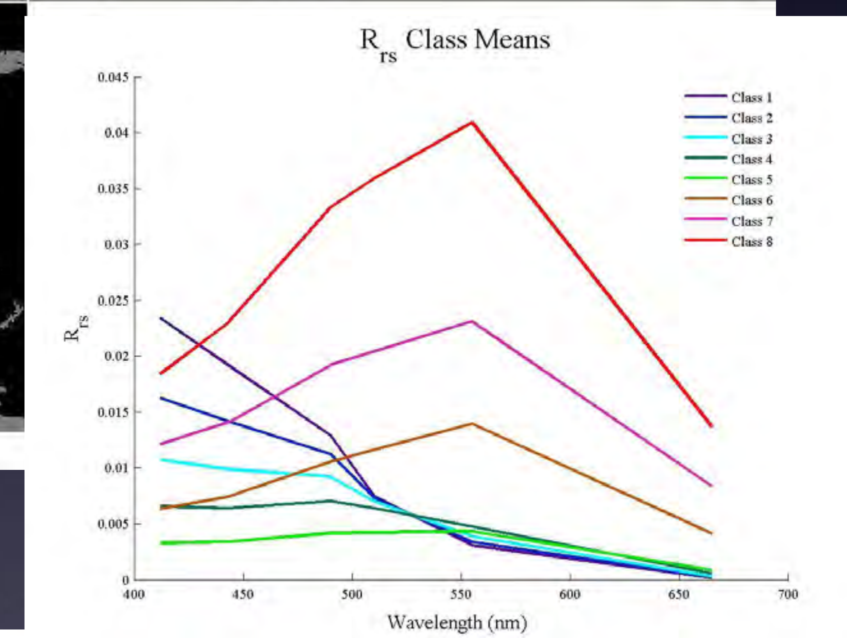
i = 1…8

Aqua OC3 Error
27
52
55
72
63
123
57
83

= Uncertainty image



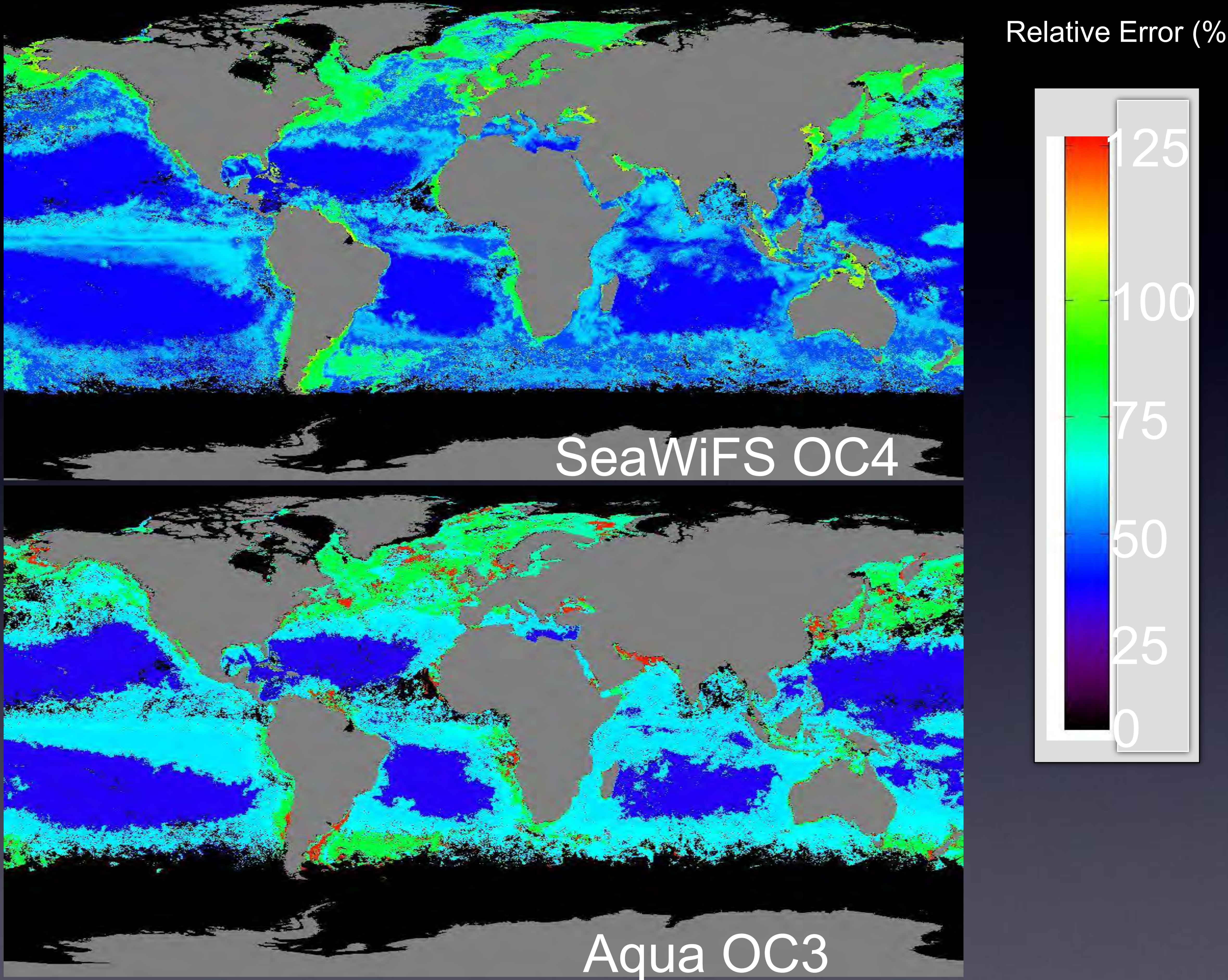


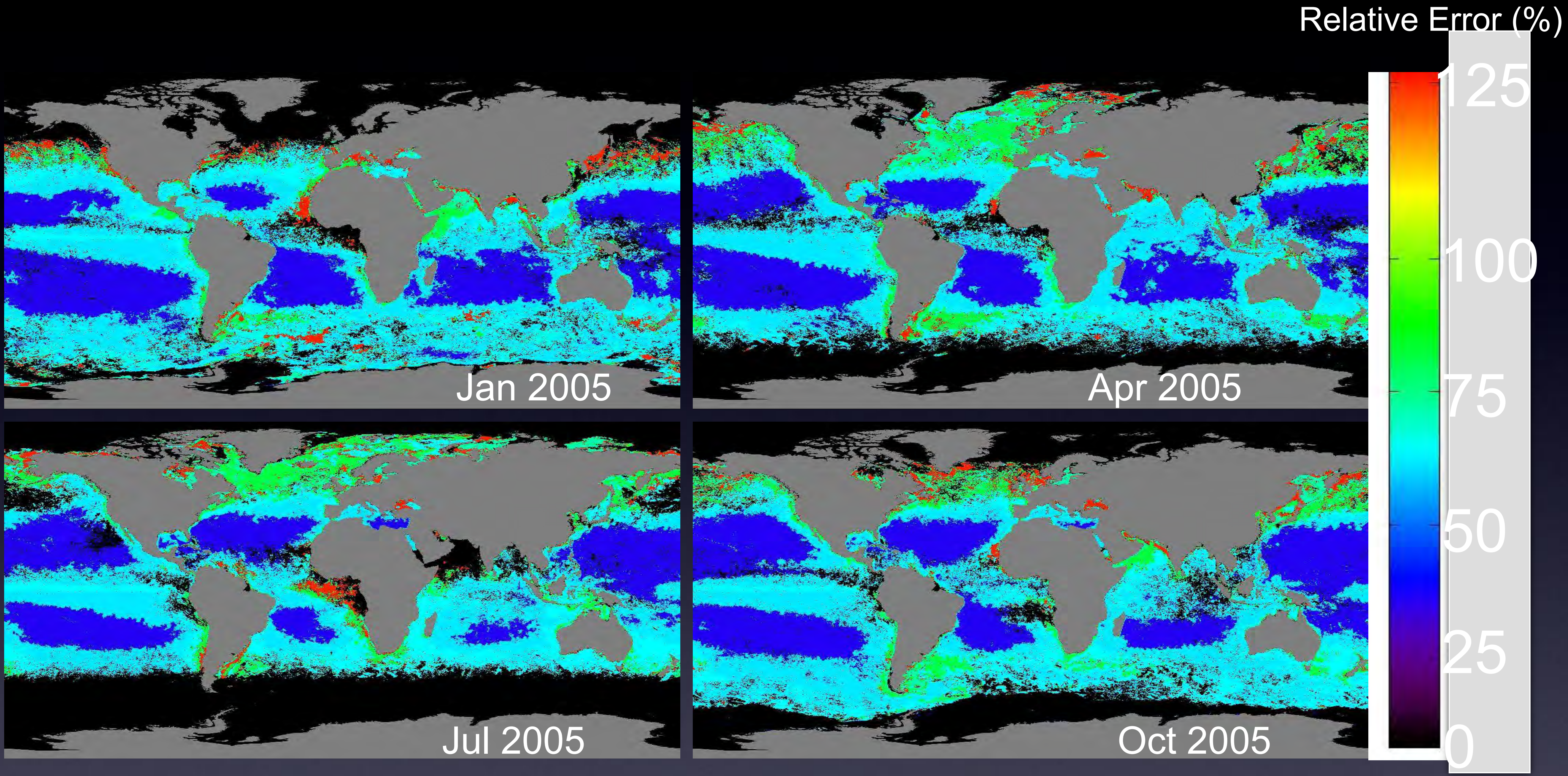


Membership

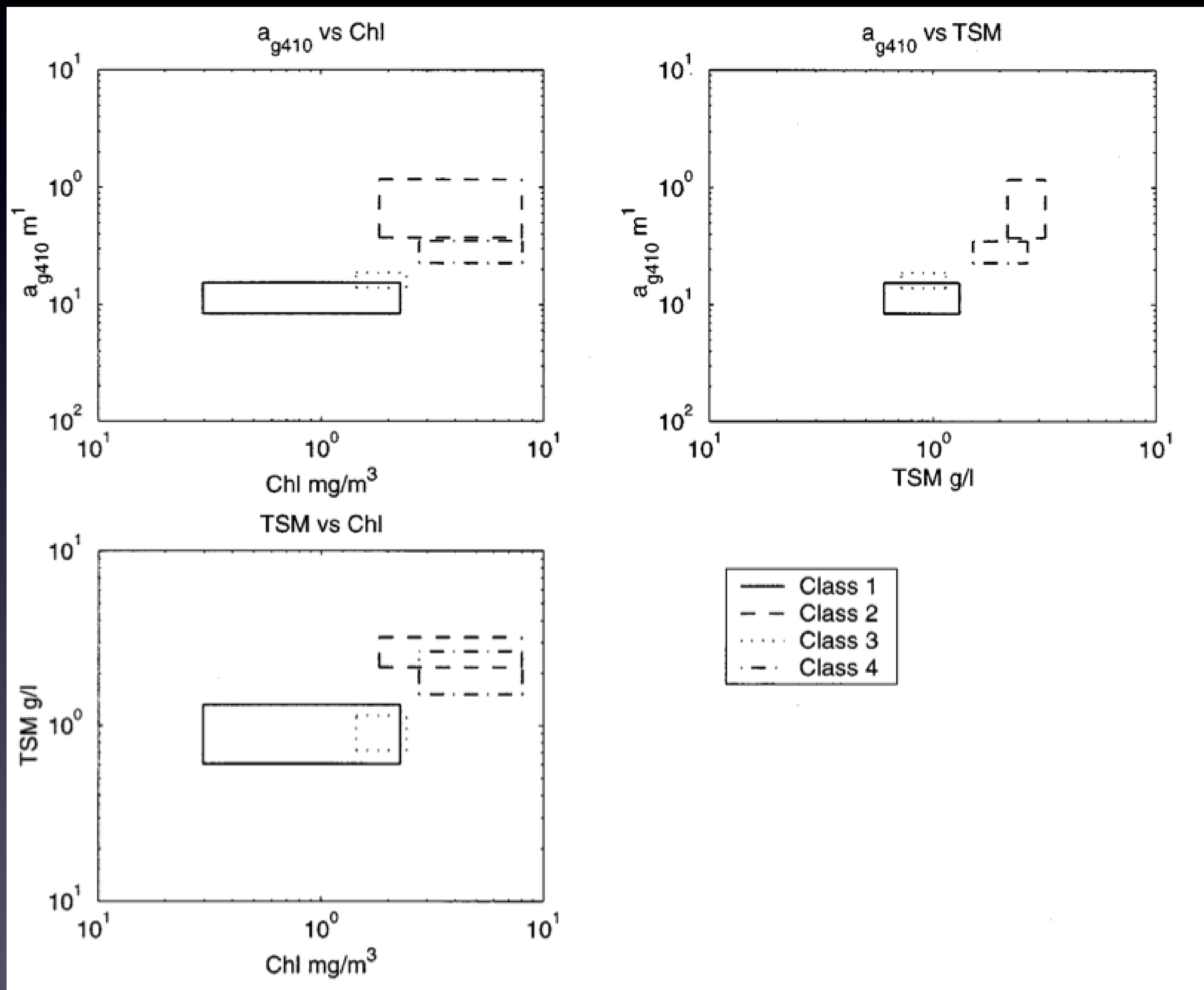
0 1

May 2005




Relative Error (%)
125
100
75
50
25
0
Jan 2005
Apr 2005
Jul 2005
Oct 2005

Bi-variate Distribution Function of Optical Constituents



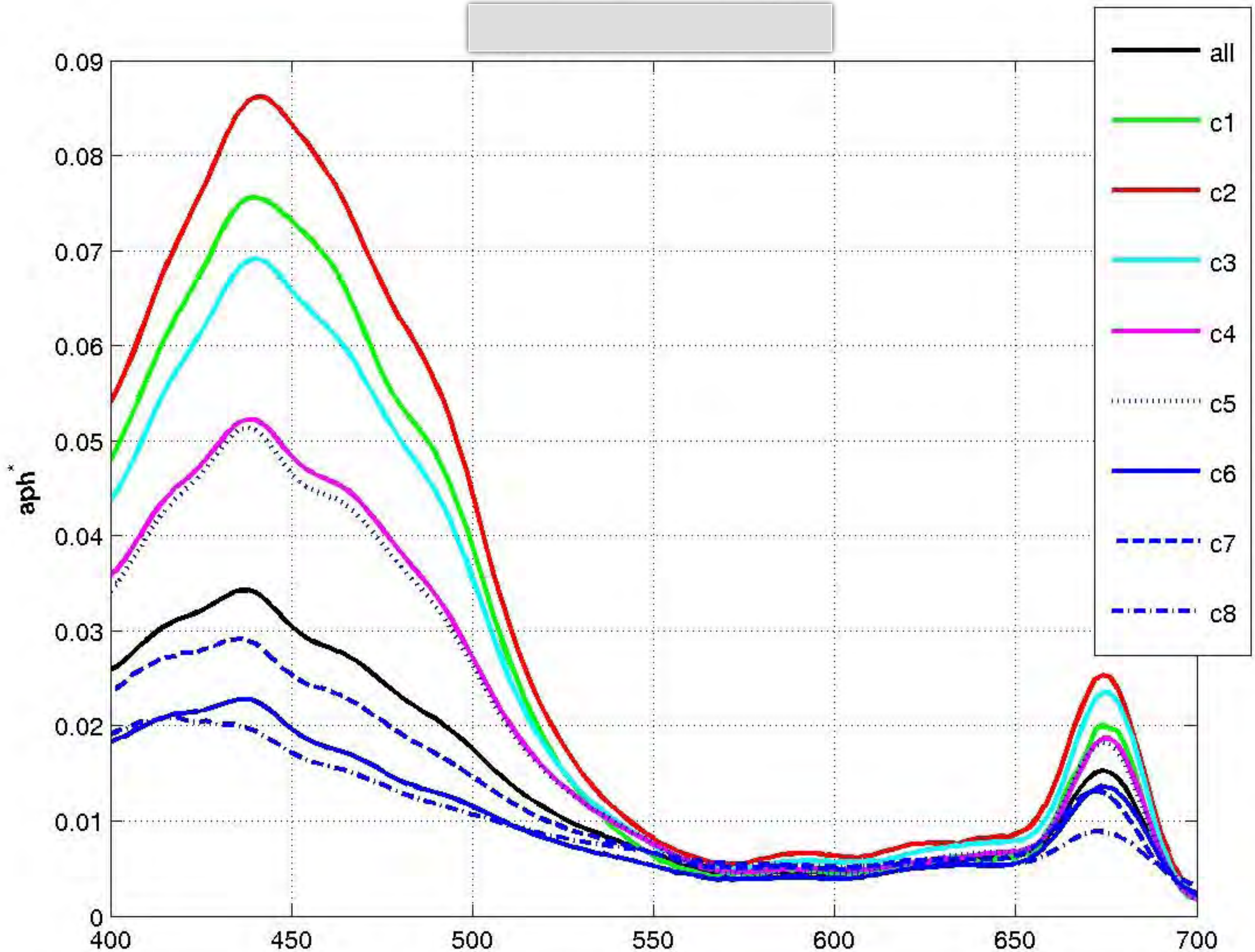
Class specific steps of QAA

Table 1
Steps of deriving phytoplankton absorption spectrum from hyperspectral remote-sensing reflectance

Steps	Property	Derivation
Step 0	r_{rs}	$= R_{rs}/(0.52 + 1.7R_{rs})$
Step 1	$u(\lambda) = \dfrac{bb(\lambda)}{a(\lambda) + bb(\lambda)}$	$= \dfrac{-0.0895 + \sqrt{(0.0895)^2 + 4g_1 r_{rs}(\lambda)}}{20.1247}$
Step 2	$a(\lambda_0)$: $a(555)$ or $a(640)$	$a(555) = 0.0596 + 0.2(a(440)_i - 0.01)$, $a(440)_i = \exp(-1.8 - 1.4\rho + 0.2\rho^2)$, $\rho = \ln(r_{rs}(440)/r_{rs}(555))$
Step 3	$bb_p(\lambda_0)$	$= \dfrac{u(\lambda_0)a(\lambda_0)}{1 - u(\lambda_0)} - bb_w(\lambda_0)$
Step 4	Y	$= 2.2(1 - 1.2e^{-0.9 r_{rs}(440)/r_{rs}(555)})$
Step 5	$bb_p(\lambda)$	$= bb_p(\lambda_0)\left(\dfrac{\lambda_0}{\lambda}\right)^Y$
Step 6	$a(\lambda)$	$= \dfrac{(1 - u(\lambda))(bb_w(\lambda) + bb_p(\lambda))}{u(\lambda)}$
Step 7	$\zeta = a_\phi(410)/a_\phi(440)$	$= 0.71 + \dfrac{0.06}{0.8 + r_{rs}(440)/r_{rs}(555)}$
Step 8	$\xi = a_g(410)/a_g(440)$	$= e^{-S(410-440)}$
Step 9	$a_g(440)$	$= \dfrac{(a(410) - \zeta a(440))}{\xi - \zeta} - \dfrac{(a_w(410) - \zeta a_w(440))}{\xi - \zeta}$
Step 10	$a_\phi(\lambda)$	$= a(\lambda) - a_g(440)e^{-S(\lambda-440)} - a_w(\lambda)$

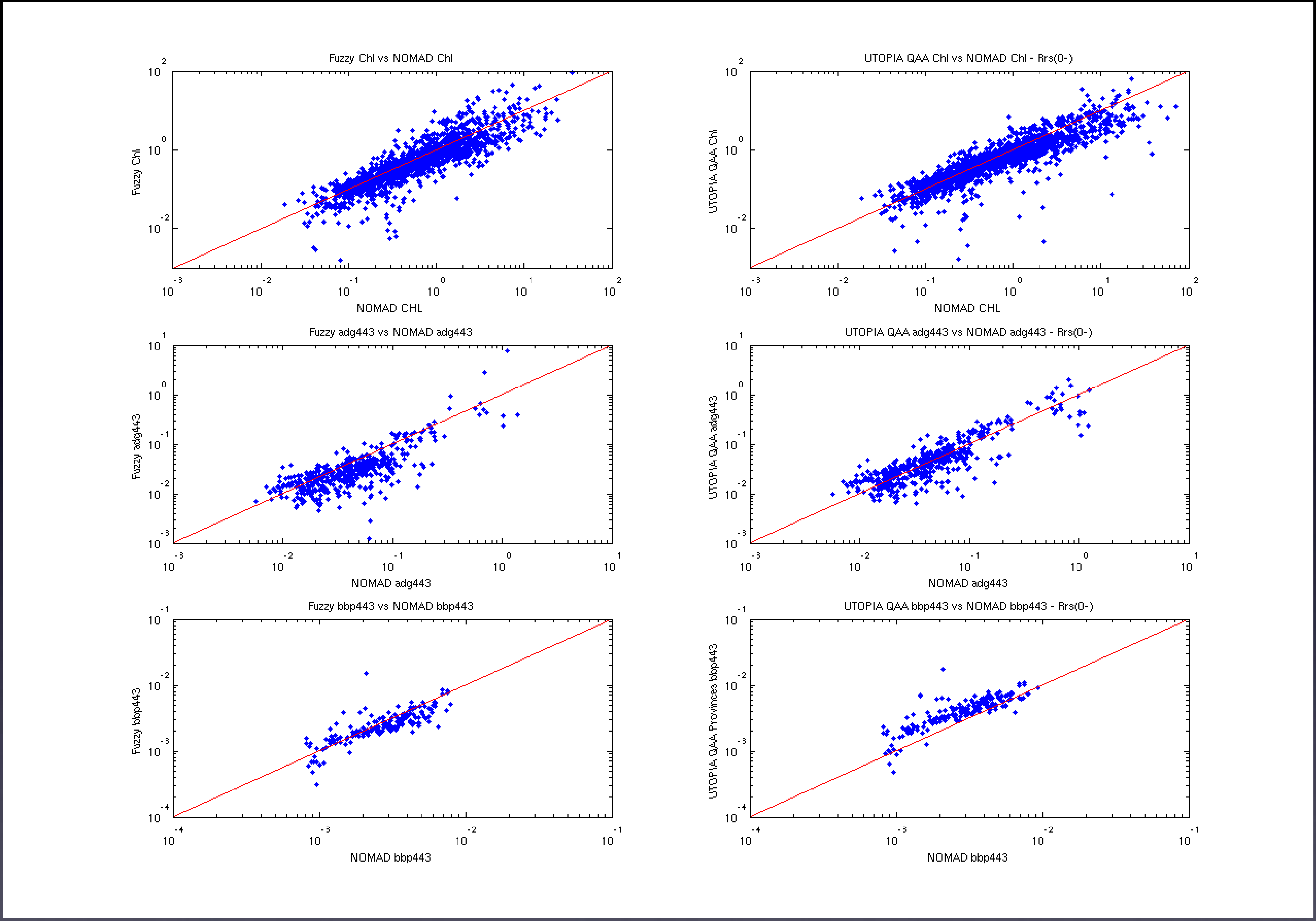
Class –based GSM	Class –based QAA
S_{gd} varies based on class [0.0175,0.0164,0.0139,0.0147, 0.0153,0.0128,0.0138,0.0121] $a_{ph}^*(\lambda)$ varies dependent on class η (i.e. slope of bbp) using Carder's relationship	S_{gd} variable based on class $a_t(443)$ versus $r_{rs}(443)/r_{rs}(555)$ class based $a_t(555)$ versus $a_t(443)$ class based $a_{ph}(443)$ versus Chl class based $a_{ph}^*(443)$

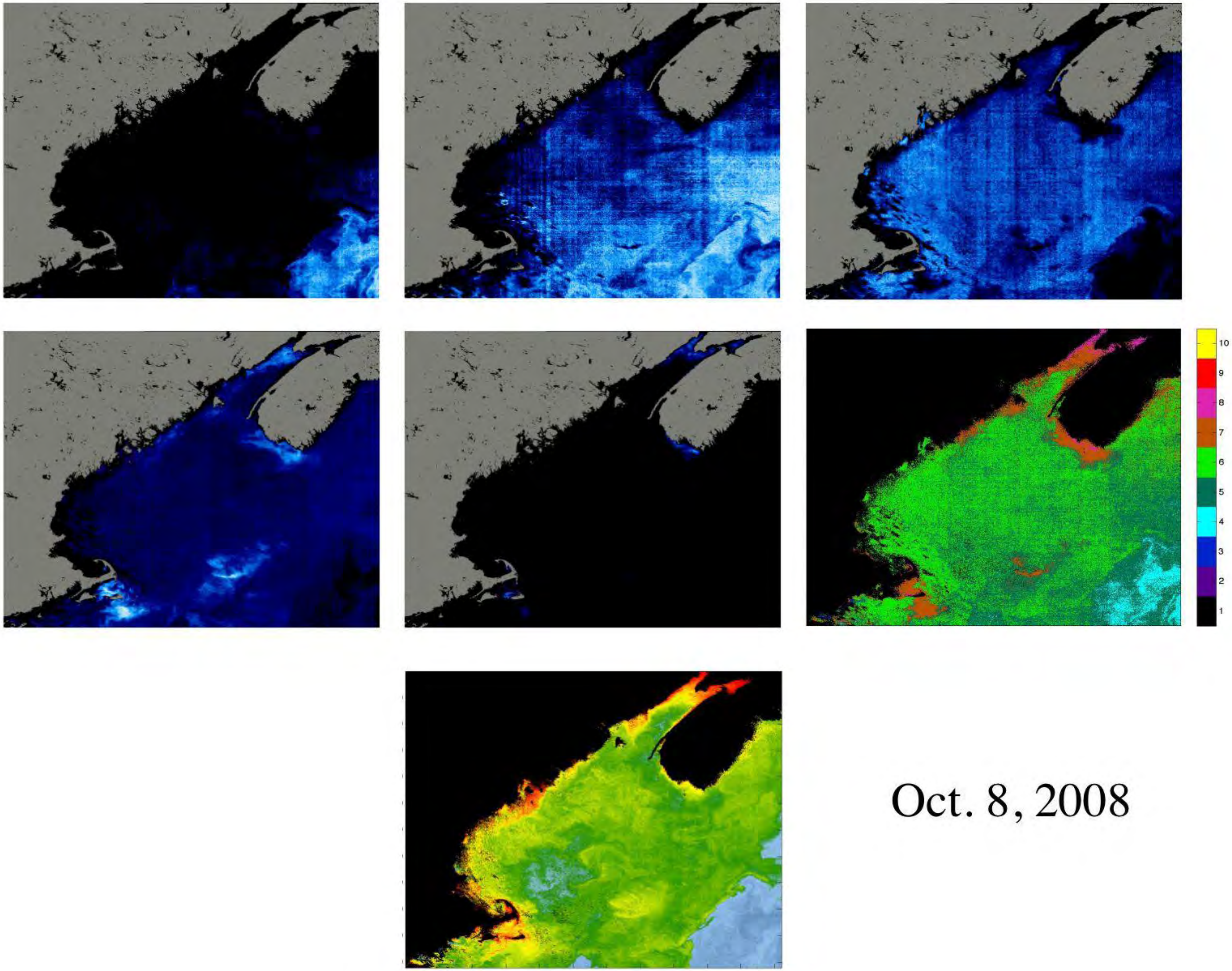
One could imagine applying a tuning algorithm (e.g. simulated annealing) to each class to determine optimal class based model coefficients.



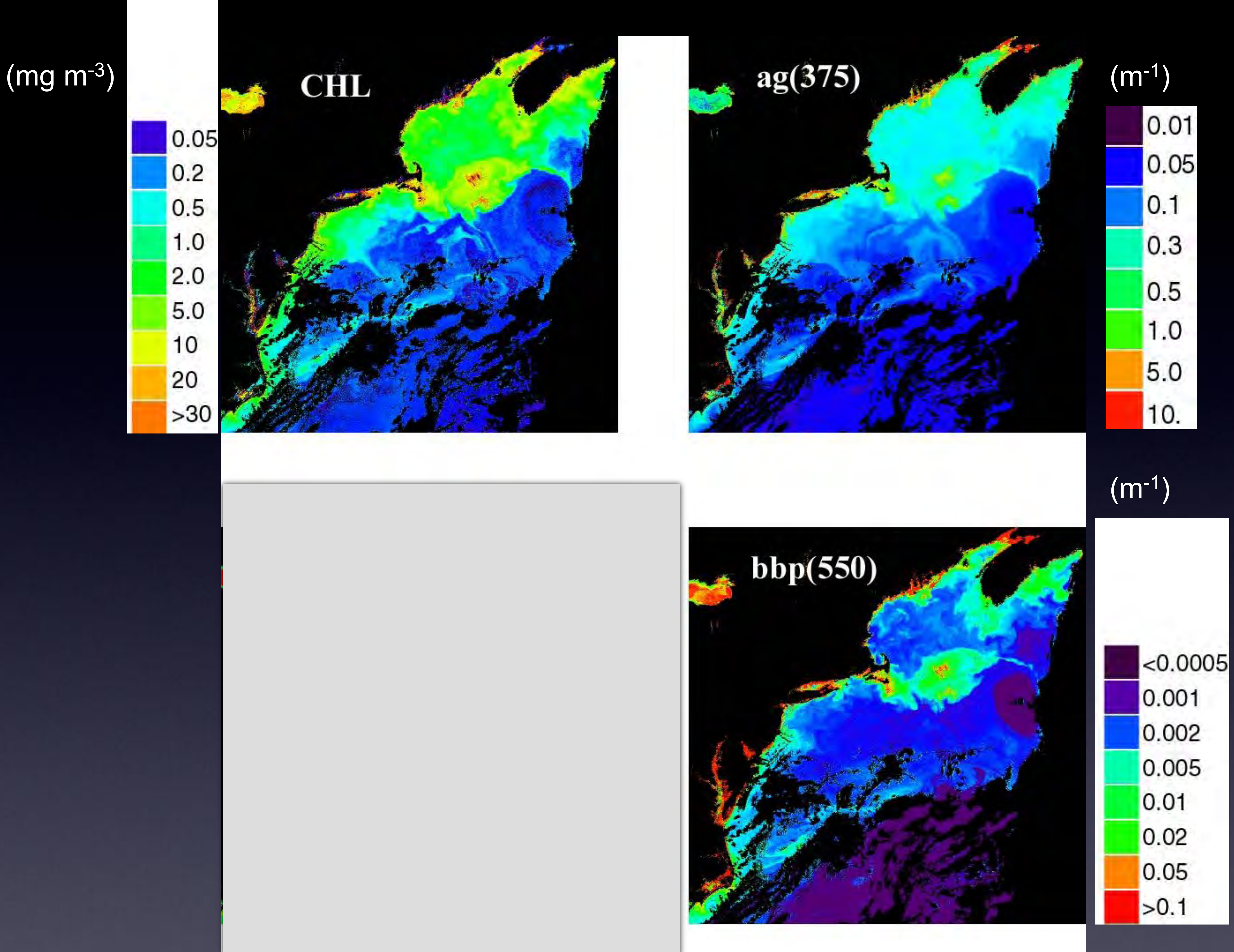
aph*
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
400
450
500
550
600
650
700
all
c1
c2
c3
c4
c5
c6
c7
c8

Amoeba - NLO   Spectral Unmixing





Oct. 8, 2008



September 12

Conclusions

- Fuzzy logic based dynamic provinces provide a powerful tool for describing the optical variability of the world oceans.
- Effective in identifying bio-optical “end-members” locations for use in identifying cal/val sites, as well as identifying “under-sampled” optical water types.
- Can also be used to determine spatial uncertainty estimates benefiting from the availability of the membership functions.
- Statistically rigorous means of parameterizing bio-optical models.
- Capable of describing the strong non-linearity of optical variability across many decades of variability

Uncovered, relevant topics

- Inversion in optically shallow waters
- Effect on inversion of transpectral processes (fluorescence and Raman scattering)
- Detail on ATBDs of specific space agency algorithms
- Ambiguity issues
- Data merging

Current efforts/Emphasis

- Using red NIR part of spectrum in highly turbid waters
- Systematic validation of IOP inversion
- Uncertianty estimate
- Algorithms & Data merging