

# Lectures on Optical Oceanography and Ocean Color Remote Sensing

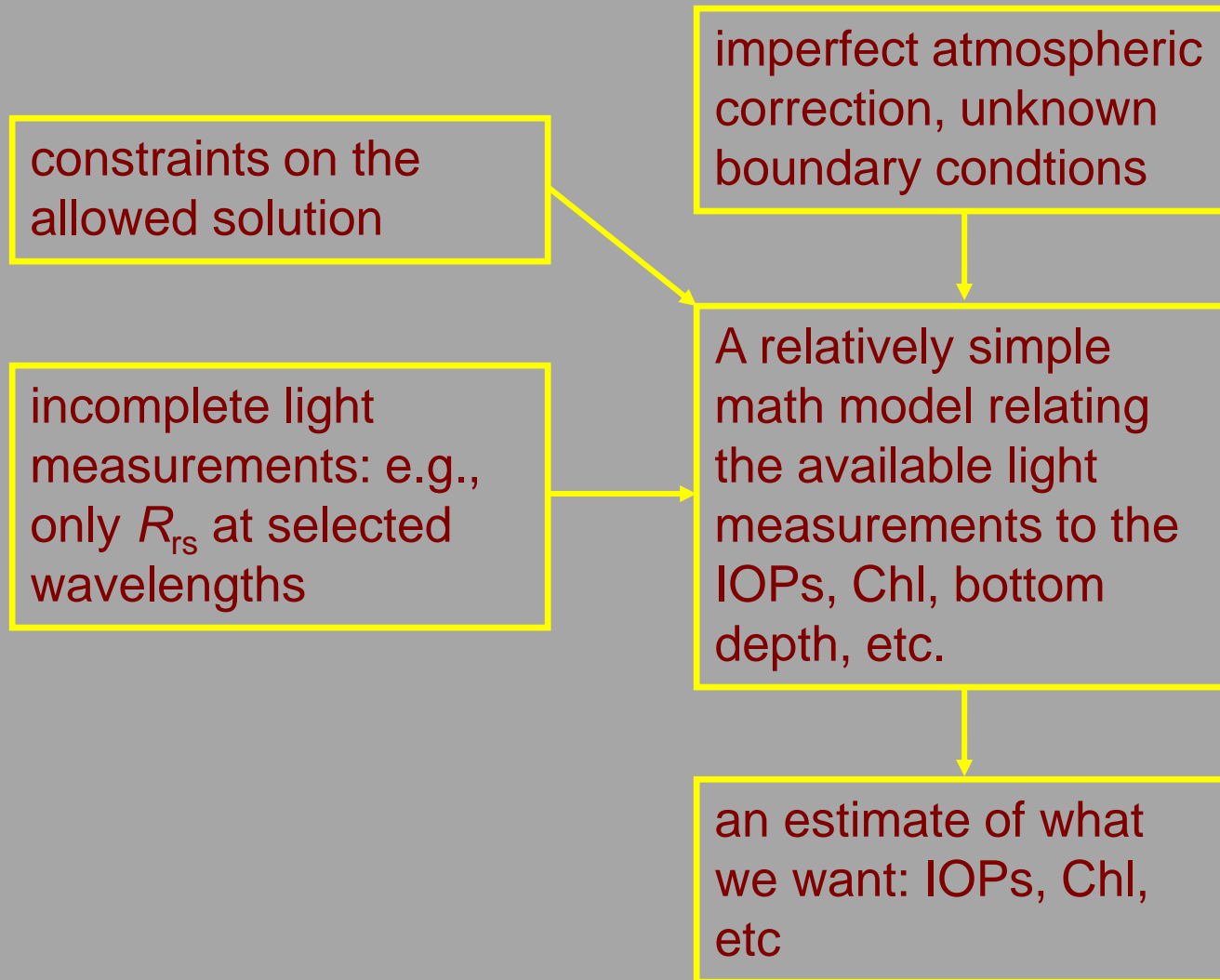
Curtis Mobley

## Statistical Methods for Remote Sensing

Schiller Coastal Studies Center

Copyright © 2021 by Curtis D. Mobley

# Recall the Remote-Sensing Inverse Problem



# Statistical Inverse Models

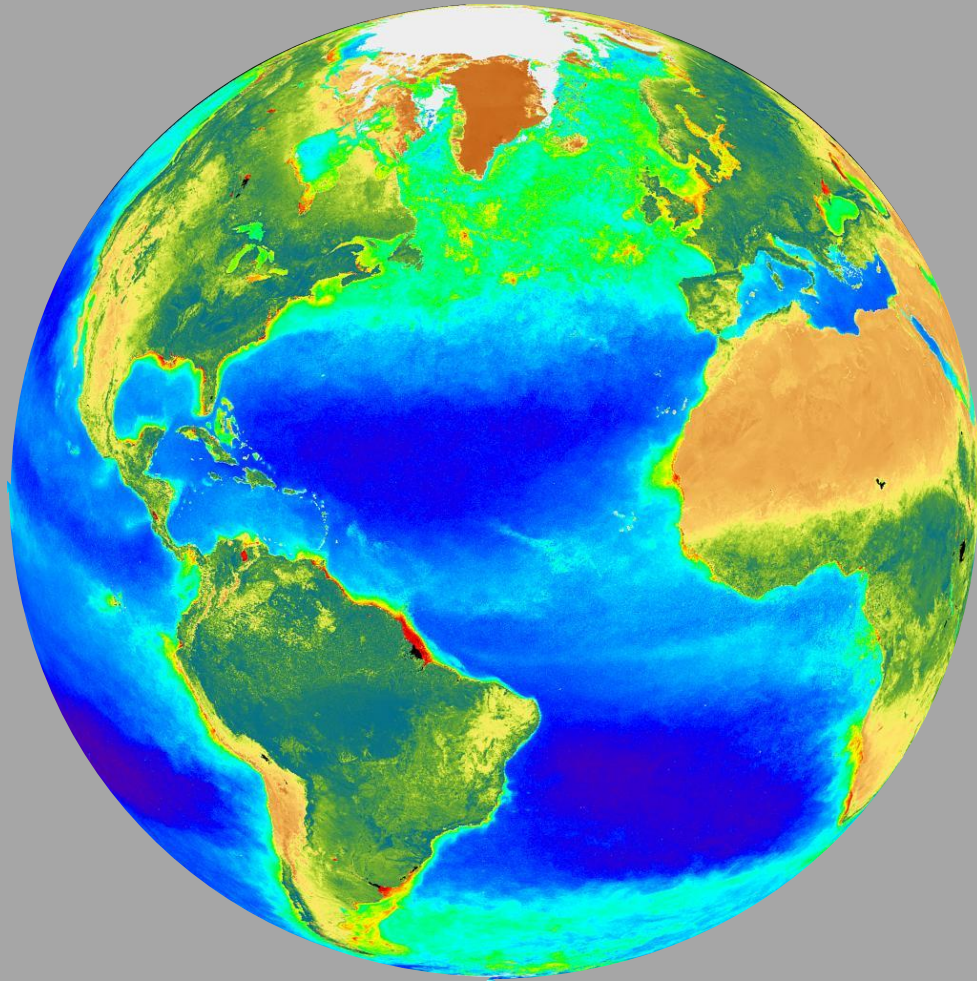
One family of simple math models relating the available measurements to what we want is *statistical* models.

These models are essentially just correlational models obtained from inspection of data sets containing both the inputs ( $R_{rs}$ ) and outputs ( $Chl$ ,  $a_{CDOM}$ , water depth, etc). The models are not necessarily based on any physical insight as to why the correlation exists.

The general forms of the models contain unknown parameters (proportionality constants, weighting functions, fitting coefficients). **The parameter values are determined by forcing the model to fit data containing both the inputs and outputs.** That is, the parameter values give the statistical best-fit of the model to the data, hence the name “statistical” or “empirical” models.

After the parameters have been determined using known inputs and outputs, the model with the same parameter values can be applied to new input data, to obtain new outputs.

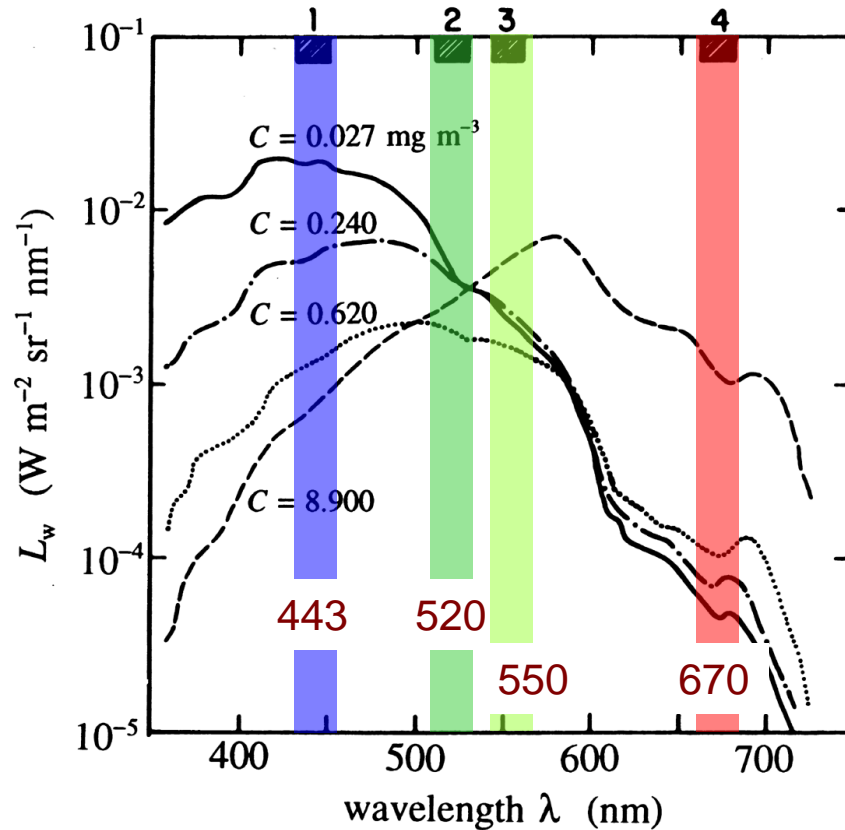
Statistical methods are how ocean color remote sensing got started 40 years ago



Two examples:

- band-ratio algorithms
- neural networks

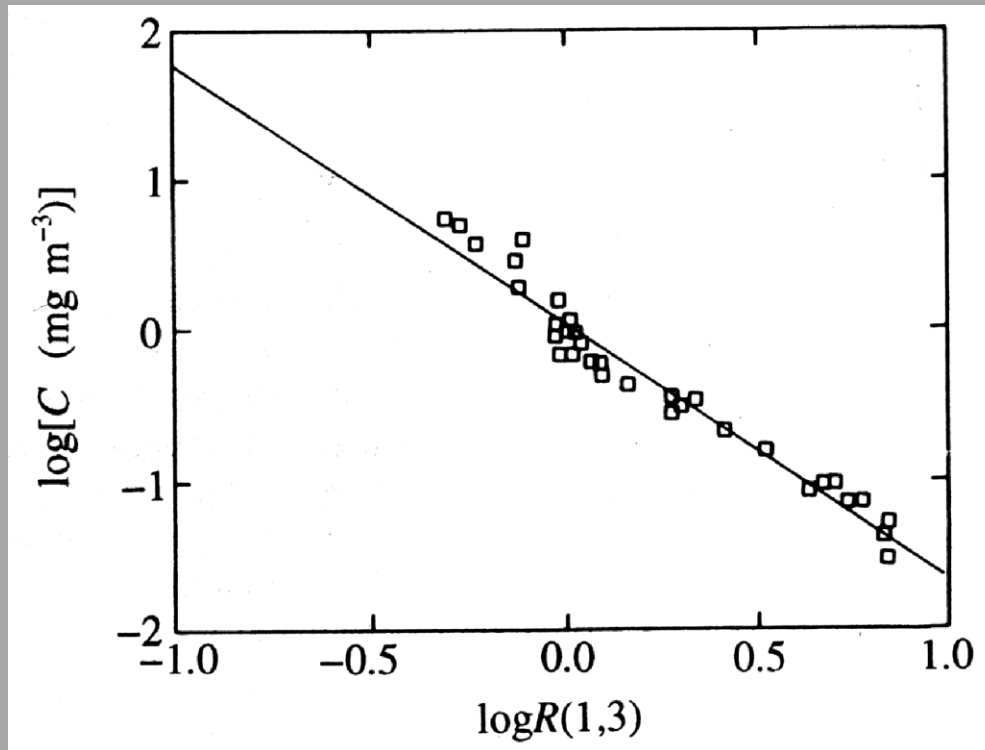
# Where It All Started



The seminal idea of ocean color remote sensing: *Chl* concentration and water-leaving radiance are correlated.

Fig. 10.1. Water-leaving radiances  $L_w$  as a function of wavelength for four chlorophyll concentrations  $C$ , in case 1 waters. The shaded regions labeled 1-4 indicate the detector bandwidths of the CZCS sensor. [redrawn from Gordon, *et al.*, (1985), by permission]

$R(1,3) = L_w(\lambda_1=443)/L_w(\lambda_3=550)$  vs *Chl*



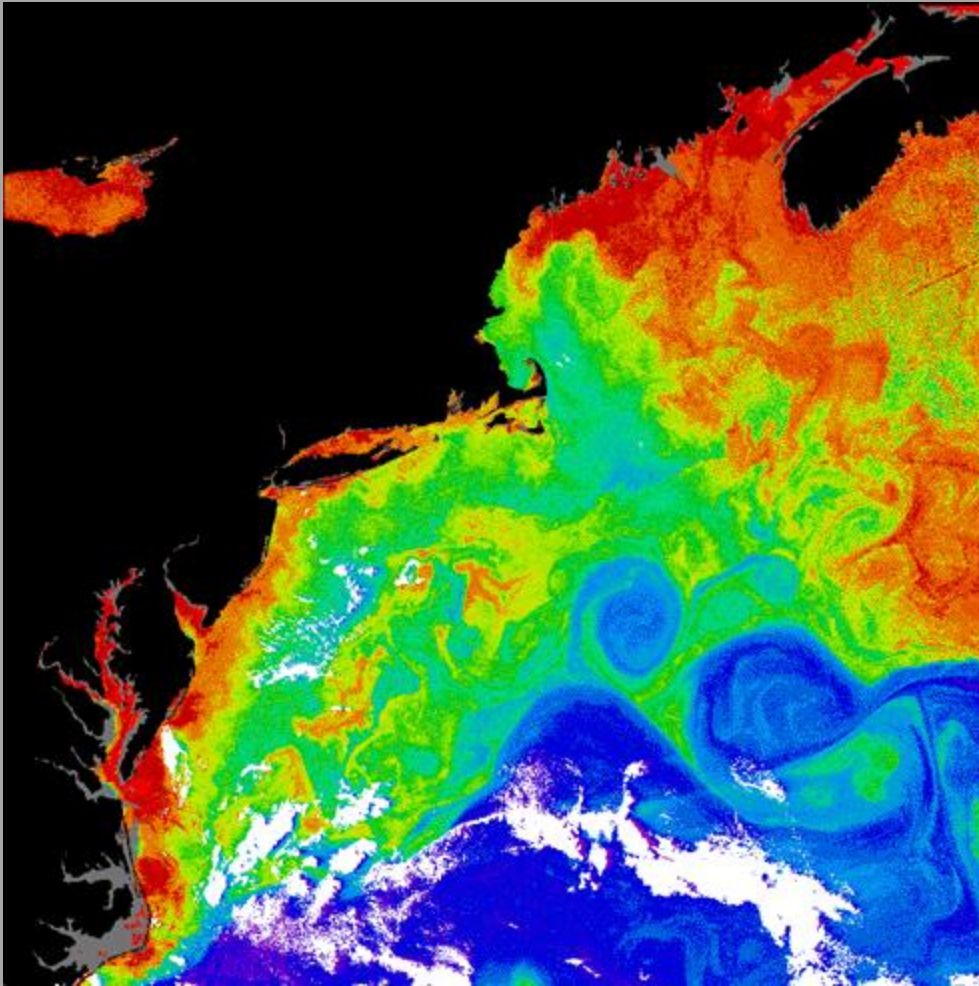
Note: only 33 data points were initially available!

This suggests the band-ratio model:

$$\log_{10}(Chl) = C_1 + C_2 \log_{10} [L_w(443)/L_w(550)]$$

$C_1$  and  $C_2$  are the unknown model parameters whose values are determined by a best fit to the data

# CZCS Image



*Chl* = 0.2 in blue to 30 in red

Coastal Zone Color  
Scanner (CZCS)

1978-1986

4 visible, 2 IR bands

66,000 images

revolutionized  
oceanography with  
very simple band  
ratio algorithms

# Examples of Recent Band-Ratio Algorithms

## SeaWiFS OC4v4 for Chl:

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

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

## MODIS for $K_d(490)$ :

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

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

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

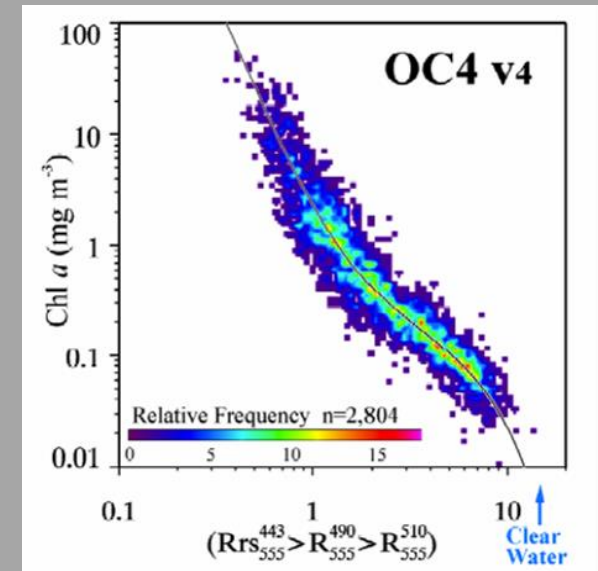
$$r_{15} = \log_{10}[R_{rs}(412)/R_{rs}(551)]$$

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

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

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

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



and so on, for dozens more....



# A Fun Project

Use HydroLight to generate some  $R_{rs}$  spectra for various Case 1 and Case 2 IOPs. Then run these  $R_{rs}$  through various band-ratio algorithms to see how the retrieved values compare with each other and with what went into HydroLight. You can find more on the [www](http://www.www).

Darieki and Stramski, *RSE*, 2004

Inadequate in-water bio-optical algorithms are one possible source of error in satellite-derived ocean color data products. Another source of error is associated with the atmospheric correction procedure, in which the water-leaving radiance is retrieved from radiance measured by a satellite sensor by subtracting the effects due to atmosphere and sea surface. Part of our database from measurements in the Baltic was used for direct comparisons with satellite-derived water-leaving radiances and other satellite-derived data products. Although our match-up data set is limited in its size, it is sufficient to reveal a consistently poor agreement between in situ-measured water-leaving radiances,  $L_{wn}(\lambda)$ , and satellite-derived  $L_{wn}(\lambda)$  from the MODIS/Terra and SeaWiFS sensors. Assuming that the in situ determinations are reasonably accurate, these match-up comparisons indicate that the current atmospheric correction for MODIS and SeaWiFS usually fails to retrieve  $L_{wn}(\lambda)$  in the Baltic. This problem is especially well pronounced in the blue spectral bands (412, 443, and 488 nm) where we observed no covariation between in situ and satellite values of  $L_{wn}(\lambda)$ .

### Acknowledgements

This research was supported by the Polish National Committee for Scientific Research—Grant PBZ-KBN 056/P04/2001. Partial support was provided by the NASA EOS Validation Grant NAG5-6466 to D. Stramski. We wish to thank all the colleagues who participated in the Baltic cruises as well as the officers and crews of the R/V *Oceania* for assistance in the collection of field data and logistical support. Special thanks are due to A. Ston, S. Kaczmarek, and P. Kowalczyk for making available the chlorophyll and CDOM data, R. Evans and W. Baringer for processing and providing MODIS data, K. Carder and R. Chen for computer code of the semianalytical algorithm, K. Kilpatrick for an update of the MODIS algorithm coefficients, and D. Clark for useful discussion.

### Appendix A

Standard MODIS and SeaWiFS in-water bio-optical algorithms examined in this study

TERRA/MODIS product number MOD 19, parameter number 13 CZCS total pigment concentration—CZCS\_pigm (Clark, 1997; K. Kilpatrick, private communication, April 2002).

$CZCS\_pigm = 10^{(aX^2 + bX^2 + cX + d)/e}$   
 $X = \log_{10} [L_{wn}(443)/L_{wn}(551)]$   
where the coefficients for the high  $X$  are:  
 $a = -1.4443$ ,  $b = 1.4947$ ,  $c = -1.5283$ ,  $d = -0.0433$ , and  $e = 1$ ,  
and for the low  $X$ :  
 $a = -5.0511$ ,  $b = 2.8952$ ,  $c = -0.5069$ ,  $d = -0.1126$ , and  $e = 1$ ,  
and the switch point between the low and high  $X$  is 0.7368

CZCS Chl

TERRA/MODIS product number MOD 19, parameter number 14 Chlorophyll a concentration for Case 1 water—chl<sub>oc</sub>\_MODIS (Clark, 1997; K. Kilpatrick, private communication, April 2002).

$chl_{oc\_MODIS} = 10^{(aX^2 + bX + c)/d}$   
 $X = \log_{10} [(L_{wn}(443) + L_{wn}(488))/L_{wn}(551)]$   
where the coefficients for the high  $X$  are:  
 $a = -2.823$ ,  $b = 1.4947$ ,  $c = -1.5283$ ,  $d = 9.110$ ,  $d = 0.8904$ , and  $e = 1$ ,  
and for the low  $X$ :  
 $a = -8.1067$ ,  $b = 12.0707$ ,  $c = -6.0171$ ,  $d = 0.8791$ , and  $e = 1$ ,  
and the switch point between the low and high  $X$  is 0.9866

MODIS Chl Case 1

TERRA/MODIS product number MOD 23, parameter number 23 Diffuse attenuation coefficient for 490 nm—K<sub>a</sub>\_490 (Clark, 1997; K. Kilpatrick, private communication, April 2002).

$K_{a\_490} = 0.11X^2 + 0.17544X + 0.1891$   
where  $X = \log_{10} [R_{rs}(490)/R_{rs}(551)]$

MODIS K<sub>a</sub>(490)

TERRA/MODIS product number MOD 21, parameter number 26 Chlorophyll a concentration for Case 2 water (SeaWiFS Case 2) (O'Reilly et al., 2000).

$chl_{oc\_a\_2} = 10^{(0.2830 - 2.753X + 1.457X^2 + 0.658X^3 - 1.403X^4)}$   
 $X = \log_{10} [(r_{25} - 0.001) / (r_{25} - R_{rs}(443)/R_{rs}(551)) / R_{rs}(488) / R_{rs}(551)]$

MODIS Chl Case 2

TERRA/MODIS product number MOD 21, parameter number 27 Chlorophyll a concentration for Case 2 water—chl<sub>oc</sub>\_MODIS (Carder et al., 1999; K. Kilpatrick, private communication, April 2002).

The computer code of the full semianalytical algorithm was received from K. Carder and R. Chen (private communication, 2002).

For default cases, the chlorophyll a concentration was calculated from empirical algorithms:

$chl_{oc\_a\_3} = 10^{(0.289 - 3.20X + 1.2X^2)}$   
where  $X = \log_{10} [R_{rs}(488)/R_{rs}(551)]$

The phytoplankton absorption coefficient at 412 nm,  $a_p(412)$ , and CDOM absorption at 400 nm,  $a_{CDOM}(400)$  (for  $a_{CDOM}$  coefficient\_gelb), for default cases were calculated from:

$a_p(675) = 0.328[10^{(0.19 + 1.037r_{15} - 0.407r_{15} - 3.531r_{15} + 0.022r_{15}^2 - 0.008r_{15}^3)} - 0.408]$   
 $a_{CDOM}(400) = 1.5[10^{(-1.67r_{15} + 0.001r_{15}^2 + 1.702r_{15}^3)}]$

where:  
 $r_{15} = \log[R_{rs}(412)/R_{rs}(551)]$ ,  $r_{25} = \log[R_{rs}(443)/R_{rs}(551)]$ , and  
 $r_{35} = \log[R_{rs}(488)/R_{rs}(551)]$

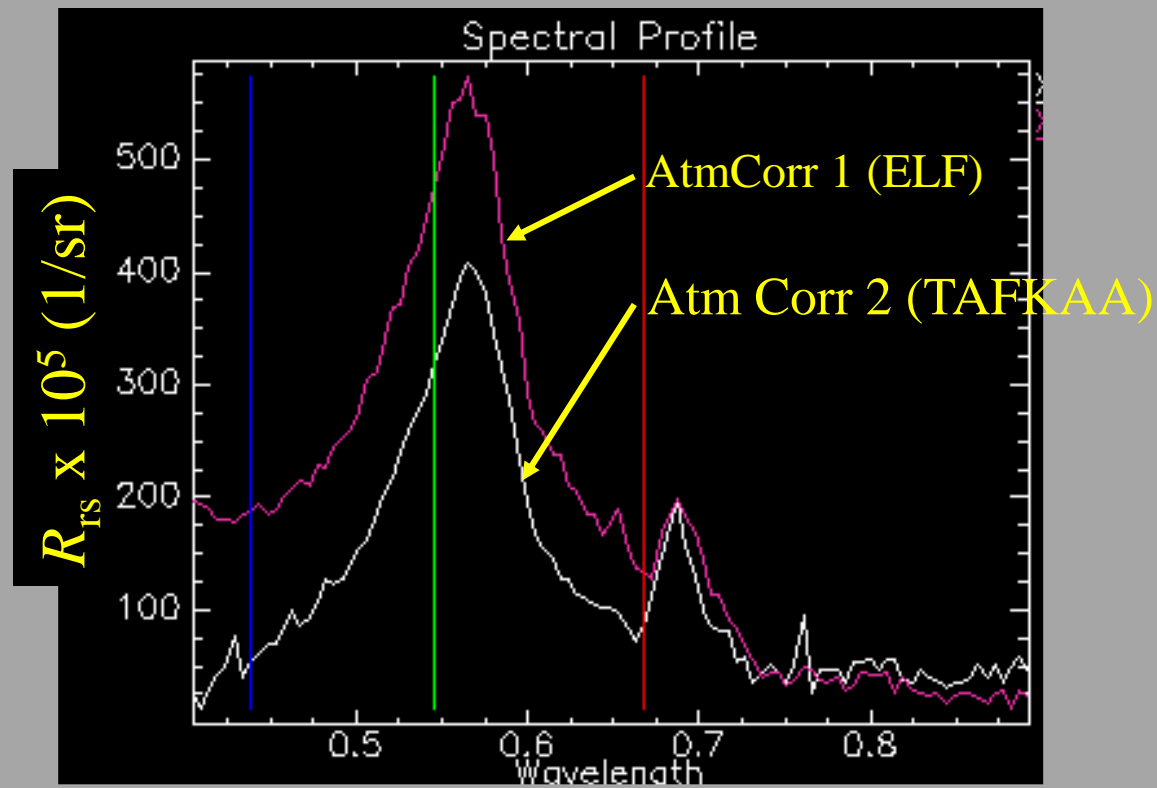
SeaWiFS OC4v4 algorithm Chlorophyll a concentration—log\_OC4v4 (O'Reilly et al., 2000).

$chl_{oc\_OC4v4} = 10^{(0.366 - 3.063X + 3.930X^2 + 0.648X^3 - 1.532X^4)}$   
 $X = \log_{10} [\max\{R_{rs}(443)/R_{rs}(551), R_{rs}(490)/R_{rs}(551), R_{rs}(510)/R_{rs}(551)\}]$

SeaWiFS Chl

# Atmospheric Correction Effects

Band-ratio algorithms can be less sensitive to bad atmospheric correction than some other techniques such as spectrum matching



# Nonuniqueness

Band-ratio algorithms are vulnerable to non-uniqueness problems because the  $R_{rs}$  ratioing throws out magnitude information that makes spectra unique. Every unique spectrum below has  $R_{rs}(490)/R_{rs}(555) = 1.71 \pm 0.01$ , which gives  $Chl = 0.59 \text{ mg/m}^3$  by the SeaWiFS OC2 algorithm; all of these spectra had  $Chl < 0.2 \text{ mg/m}^3$  (these spectra are influenced by bottom reflectance).

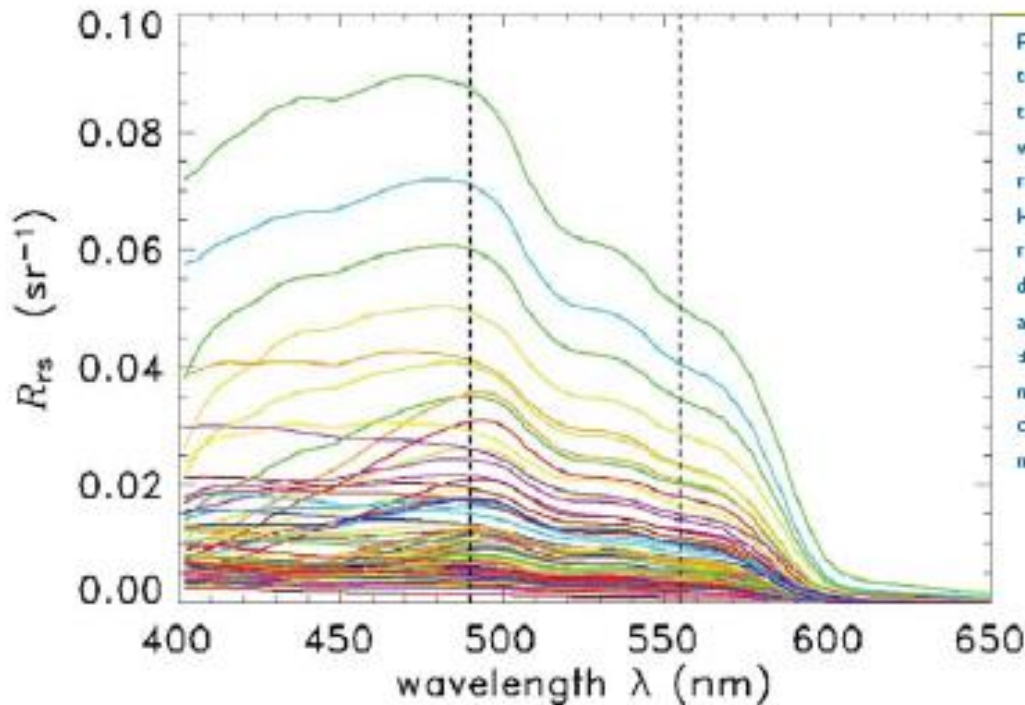
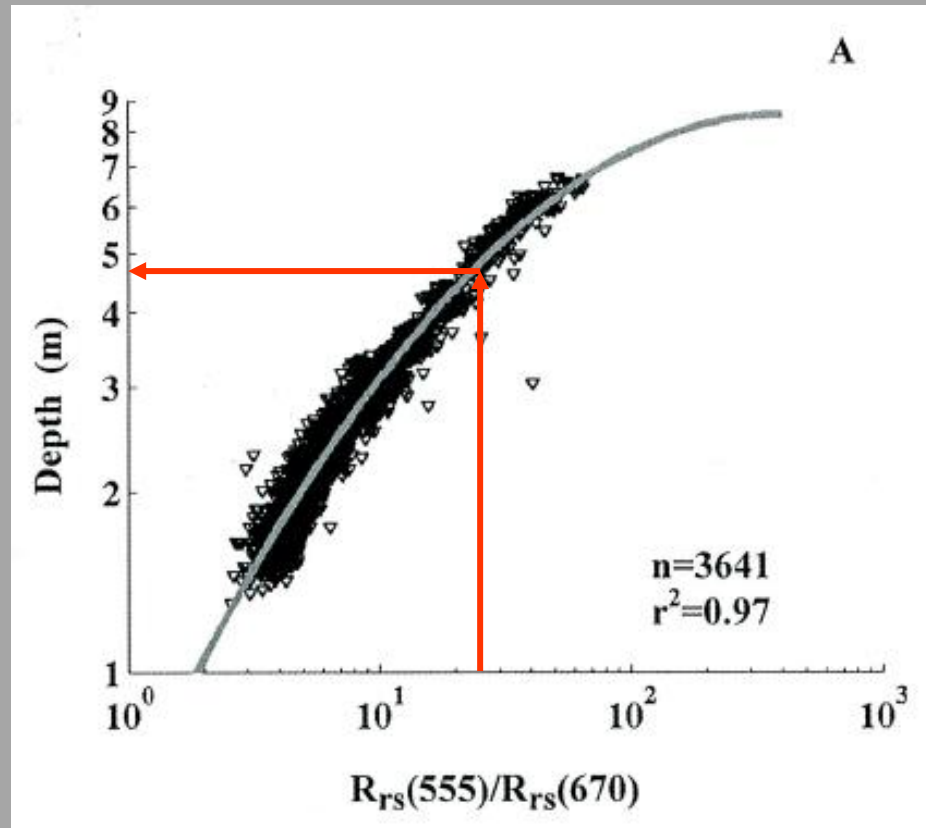


Figure 3. Chlorophyll concentration algorithms designed for multispectral instrumentation may not be useful for shallow, optically clear waters. Shown here are one hundred twenty two Hydrolight-generated remote sensing reflectance ( $R_{rs}$ ) spectra for Bahamian waters using various combinations of nine different sets of IOPs, 32 different bottom reflectances, and 22 depths between 5.5 and 50 m. These spectra are clearly unique. However, every spectrum has nearly the same remote sensing reflectance wavelength ratio:  $R_{rs}(490)/R_{rs}(555) = 1.71 \pm 0.01$  (490 and 555 nm are indicated by the vertical black dashed lines). If this ratio were applied to the commonly used SeaWiFS band-ratio algorithm (OC2; O'Reilly et al., 1998), it would give a chlorophyll concentration of  $0.59 \pm 0.01 \text{ mg Chl m}^{-3}$ . In other words, the same chlorophyll concentration would be determined for all 122 spectra despite the fact that these simulated water bodies have IOPs corresponding to chlorophyll concentrations between 0.0 (pure water) and  $0.2 \text{ mg Chl m}^{-3}$ . The OC2 algorithm fails here because of bottom effects in optically clear waters.

# Nonuniqueness

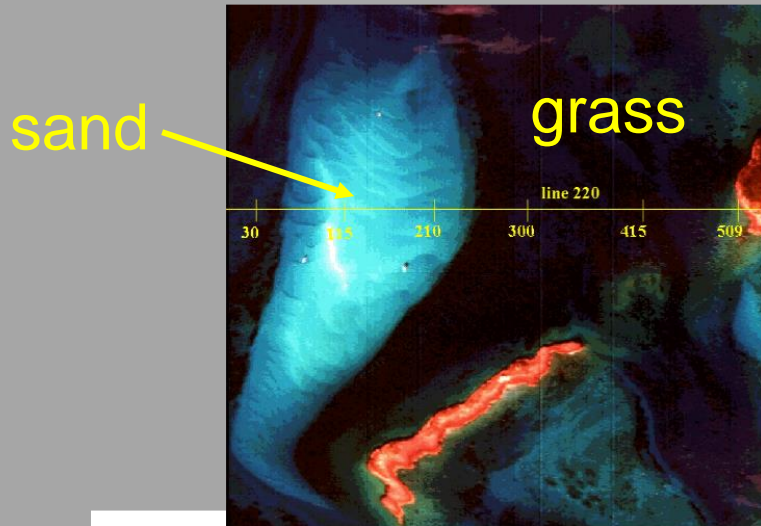
Dierssen et al. (*Limnol. Oceanogr.* 41(1), 444-455, 2003) developed a band-ratio algorithm for bottom depth in clear Bahamas waters:



$$x = \log_{10} [R_{rs}(555)/R_{rs}(670)].$$

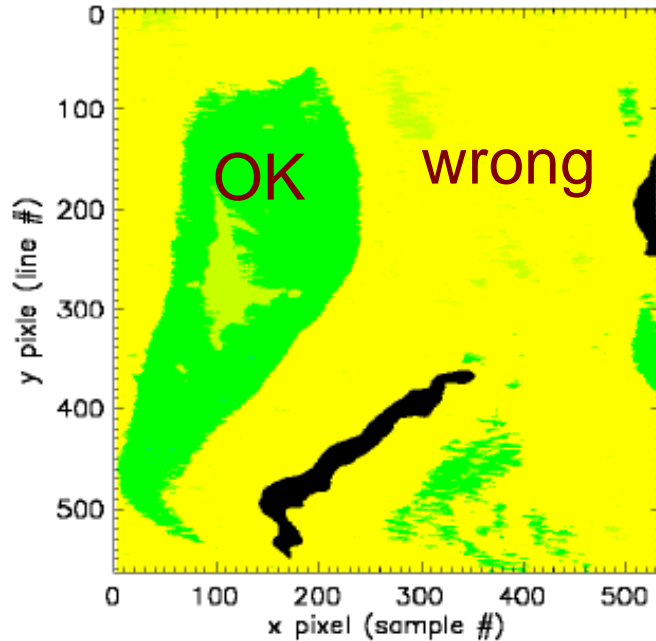
$$\log_{10} (z_b) = -0.1706 x^2 + 0.8913 x - 0.2316.$$

# Nonuniqueness

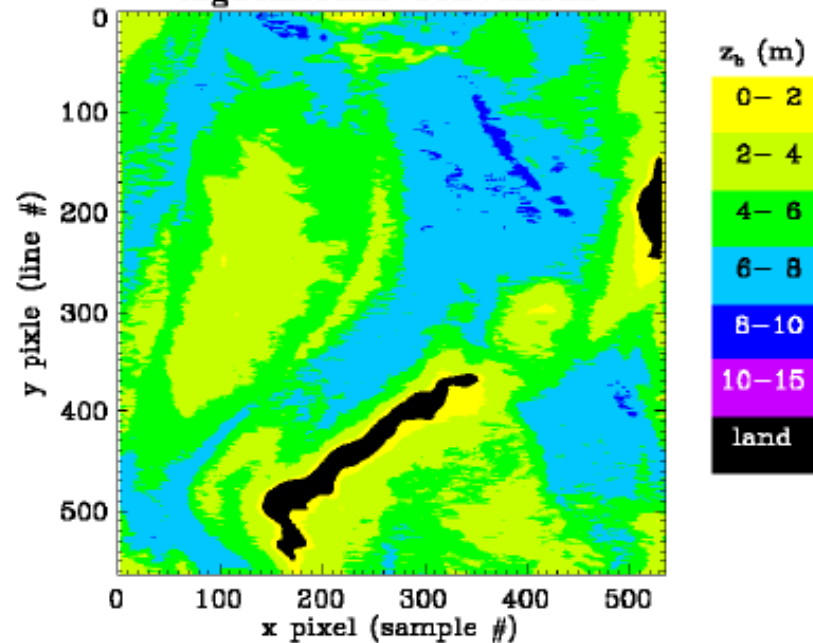


The Dierssen algorithm did OK over shallow sand bottoms, but totally failed over deeper sea grass bottoms. Why?

Dierssen et al zb model

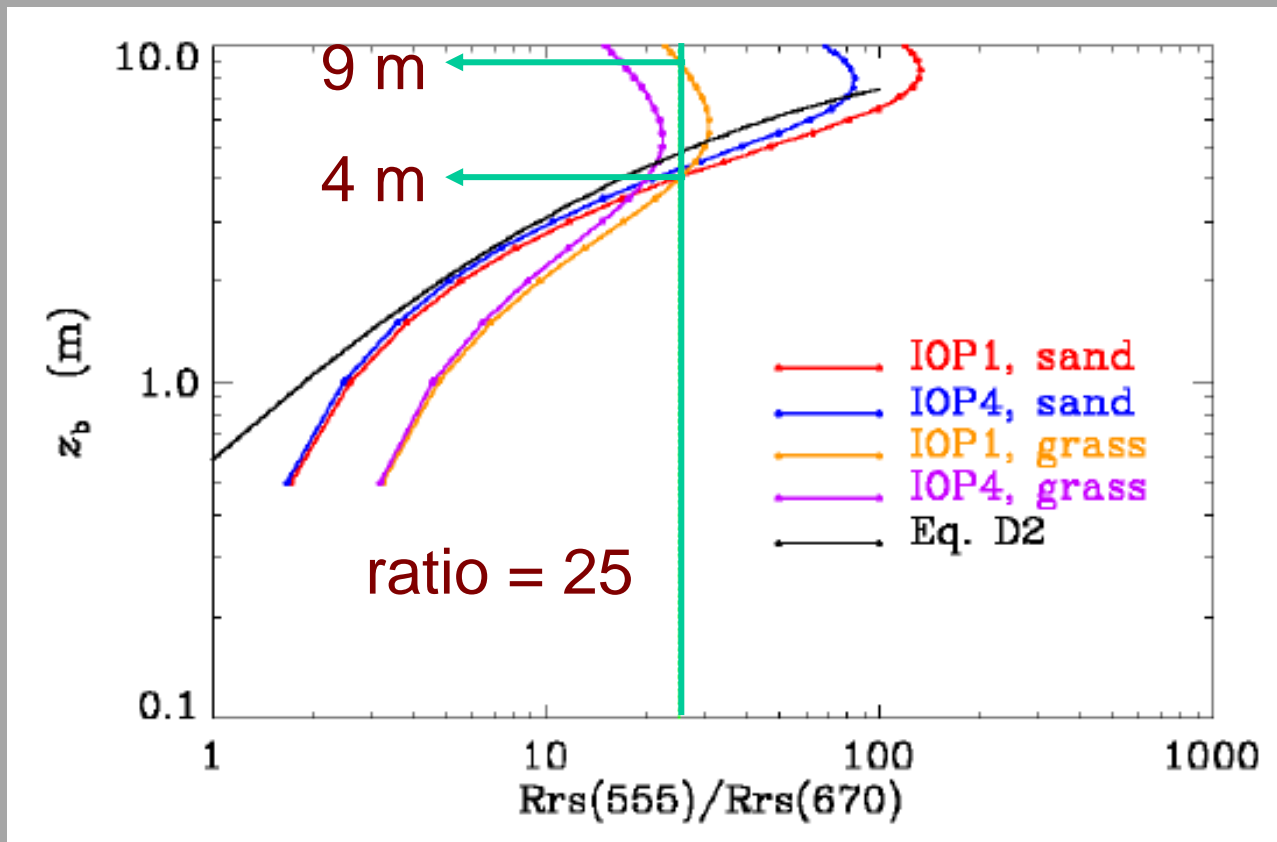


Algo1\_RGS\_5750\_w1\_UnS



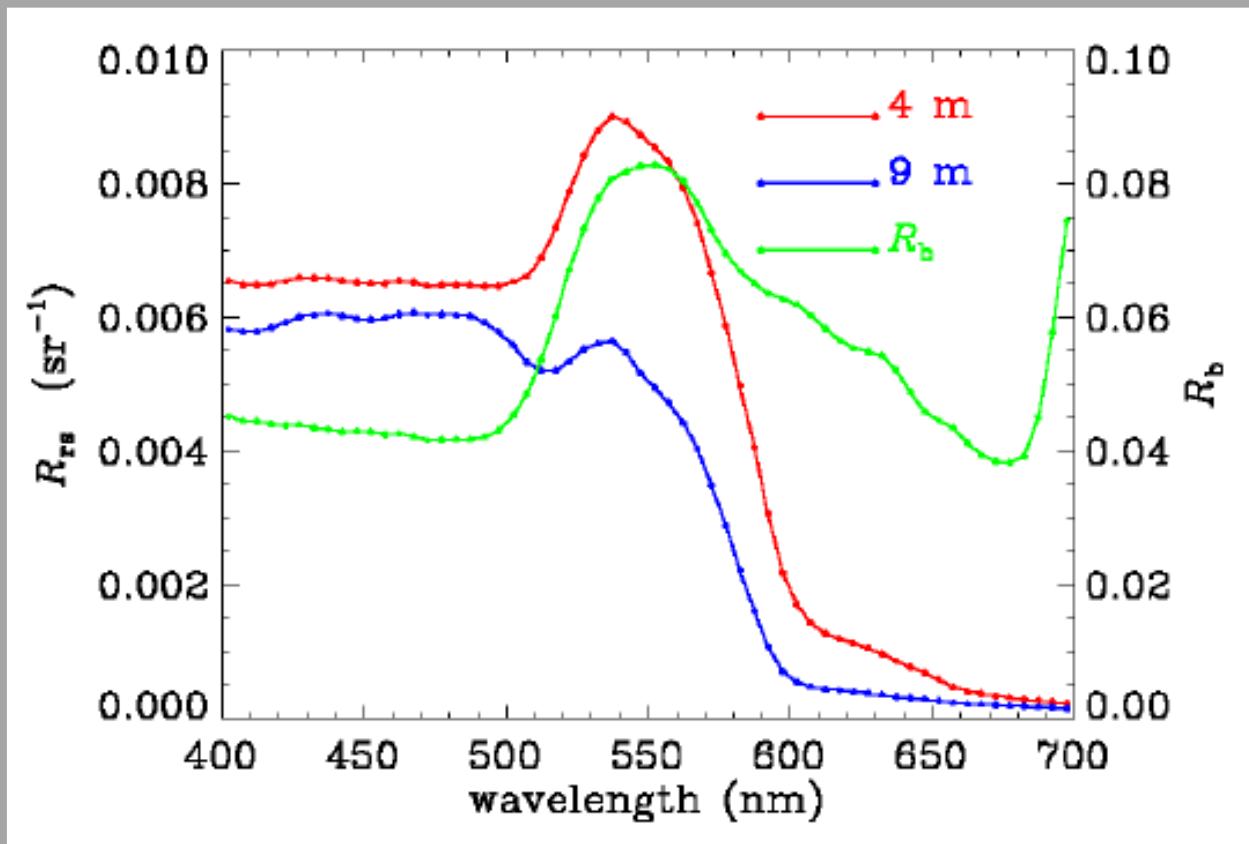
# Nonuniqueness

HydroLight simulations of  $R_{rs}(555)/R_{rs}(670)$  for two sets of IOPs and two different bottoms (sand and grass), as a function of bottom depth. Nonuniqueness for  $z_b > 5$  m and grass bottom.



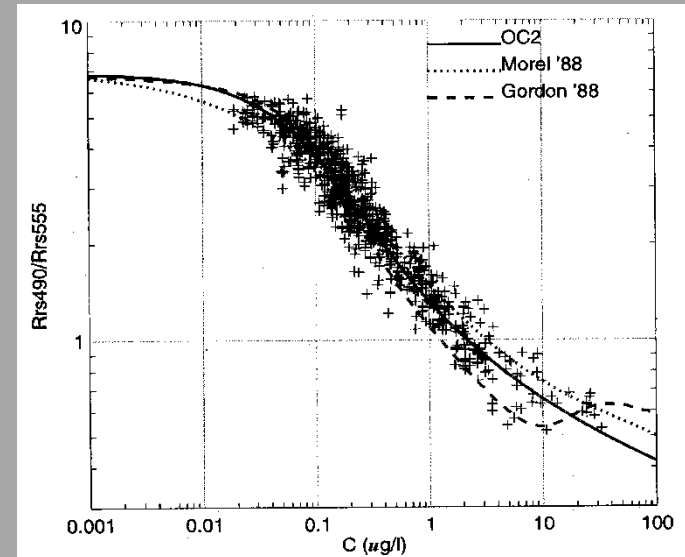
# Nonuniqueness

The  $R_{rs}$  spectra for  $z_b = 4$  and 9 m depth, grass bottom. Both spectra have  $R_{rs}(555)/R_{rs}(670) = 25 \pm 0.1$ . The Dierssen model gives  $z_b = 4.8$  m.



# Model Selection

In some situations, you can figure out (from intuition, theoretical guidance, or data analysis) the general mathematical form of the model that links the input and output (e.g., the polynomial functions that relate the band ratios to Chl). You can then use the available data (e.g., simultaneous measurements of  $R_{rs}(\lambda)$  and Chl) to get best-fit coefficients in the model via least-squares fitting.



**Figure 6.** Relationship between chlorophyll and  $R_{rs490}/R_{rs555}$  for the ocean chlorophyll 2 empirical algorithm (solid line), Morel '88 (dotted line), and Gordon '88 (dashed line).  
O'Reilly et al., JGR, 1998

**But what if you don't have any idea what the mathematical form of the model is?**



# Neural Networks

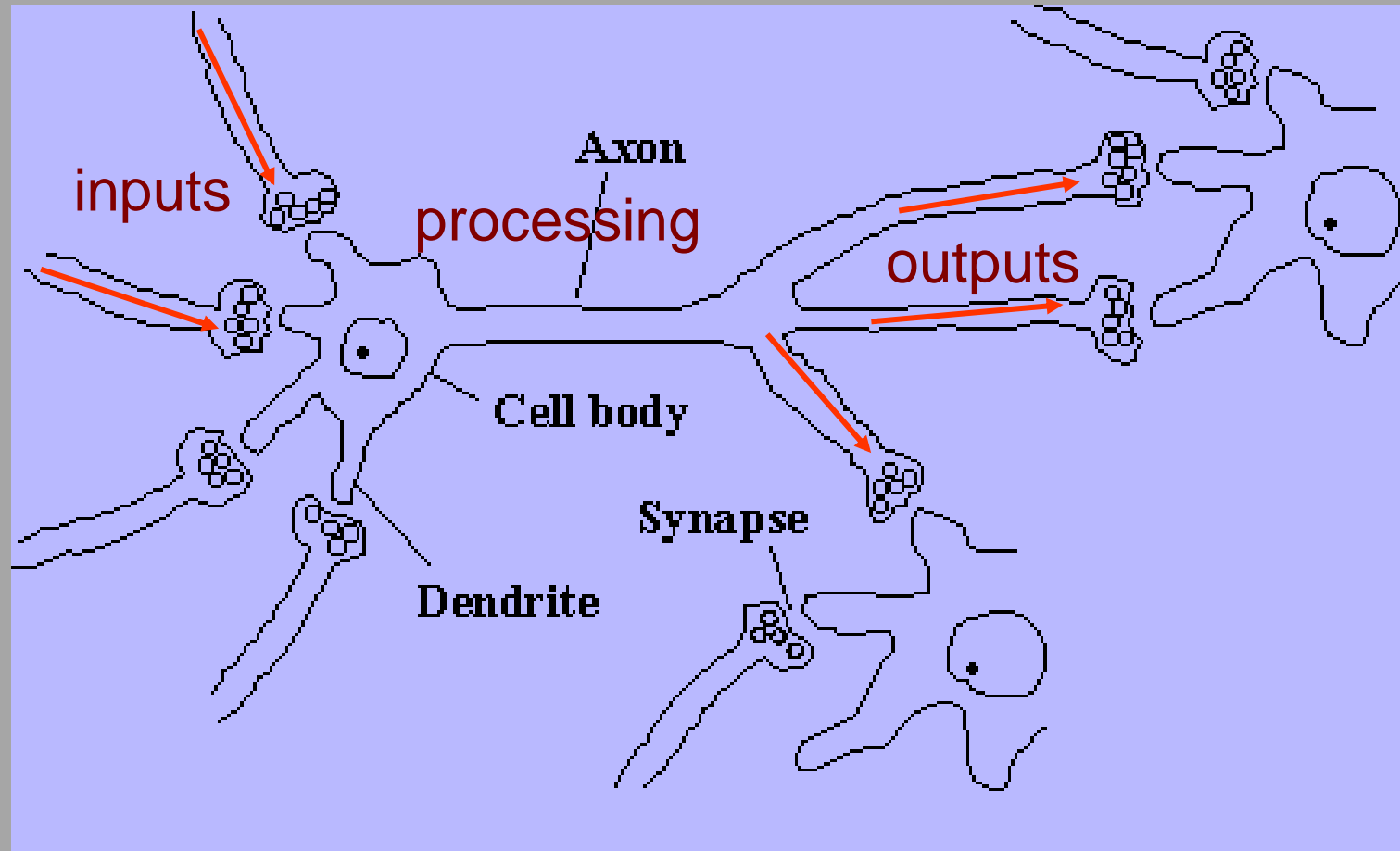
Neural networks are a form of multiprocessor computation, *based on the parallel architecture of animal brains*, with

- simple processing elements
- a high degree of connection between elements
- simple input and output (real numbers)
- adaptive interaction between elements

Neural networks are useful

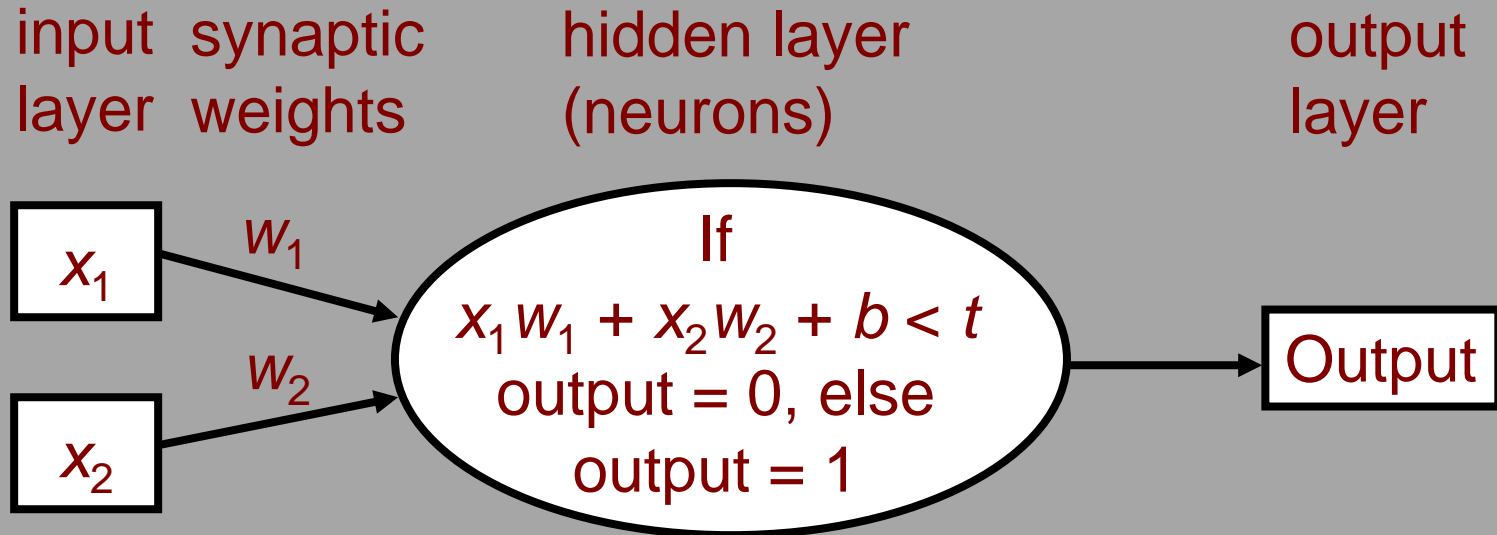
- where we don't know the mathematical form of the model linking the input and output
- where we have lots of examples of the behavior we require (lots of data to “train” the NN)
- where we need to determine the model structure from the existing data

# Biological Neural Networks



from [www.qub.ac.uk/mgt/intsys/nbiol.html](http://www.qub.ac.uk/mgt/intsys/nbiol.html)

# A Simple Artificial Neural Network

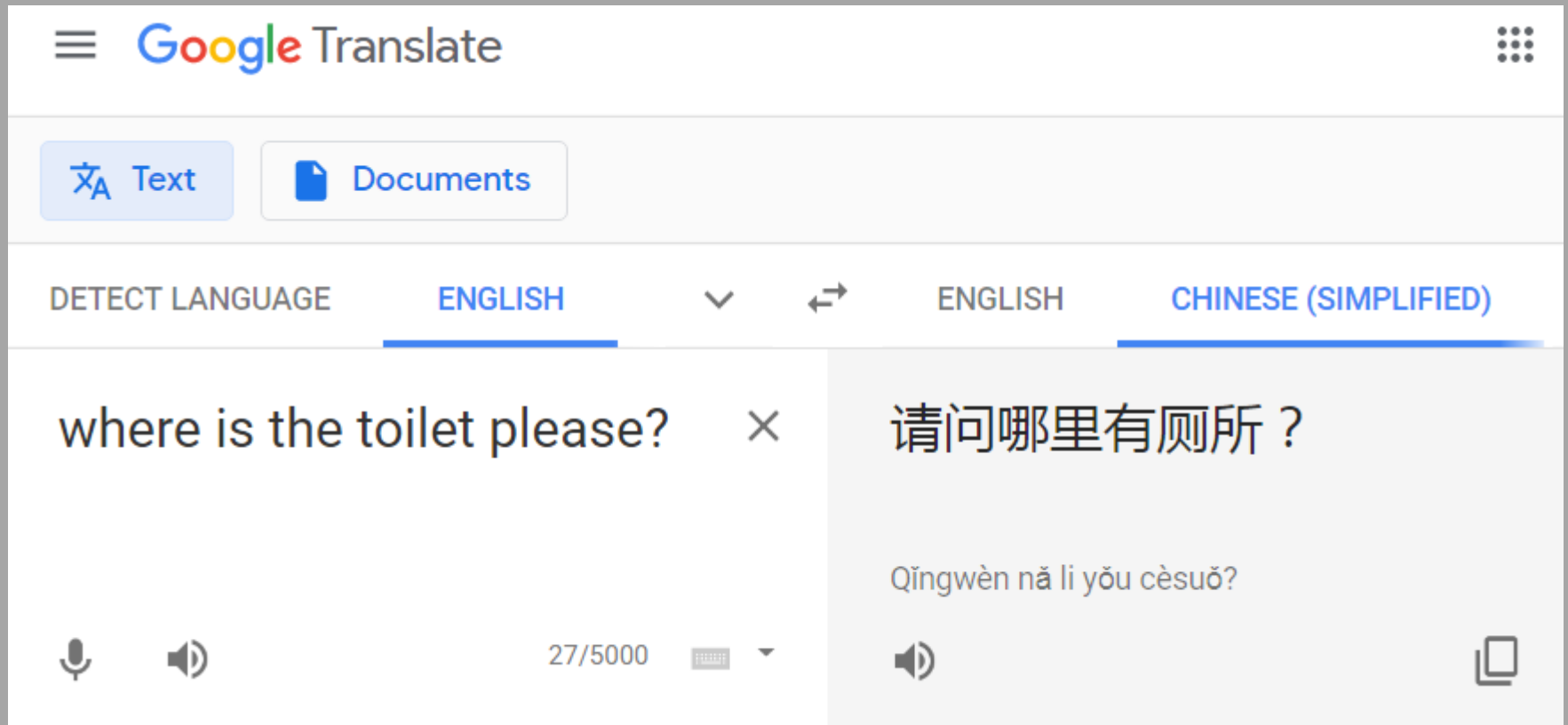


In the neuron,  $b$  is the bias,  $t$  is the threshold value

The neuron (processor) does two simple things:

- (1) it sums the weighted inputs
- (2) compares the biased sum to a threshold value to determine its output

# Example Applications of Neural Networks



Not great literature, but good enough for many purposes


# Example Applications of Neural Networks

cat photos

All Images News Videos Books More Settings Tools

About 199,000,000 results (0.60 seconds)

Images for cat



→ More images for cat

WIRED STAFF SCIENCE 06.26.12 11:15 AM

## GOOGLE'S ARTIFICIAL BRAIN LEARNS TO FIND CAT VIDEOS

BY LIAT CLARK, *Wired UK*

When computer scientists at Google's mysterious X lab built a neural network of 16,000 computer processors with one billion connections and let it browse YouTube, it did what many web users might do — it began to look for cats.

[partner id="wireduk"] The "brain" simulation was exposed to 10 million randomly selected YouTube video thumbnails over the course of three days and, after being presented with a list of 20,000 different items, it began to recognize pictures of cats using a "deep learning" algorithm. This was despite being fed no information on distinguishing features that might help identify one.

# Example Applications of Neural Networks

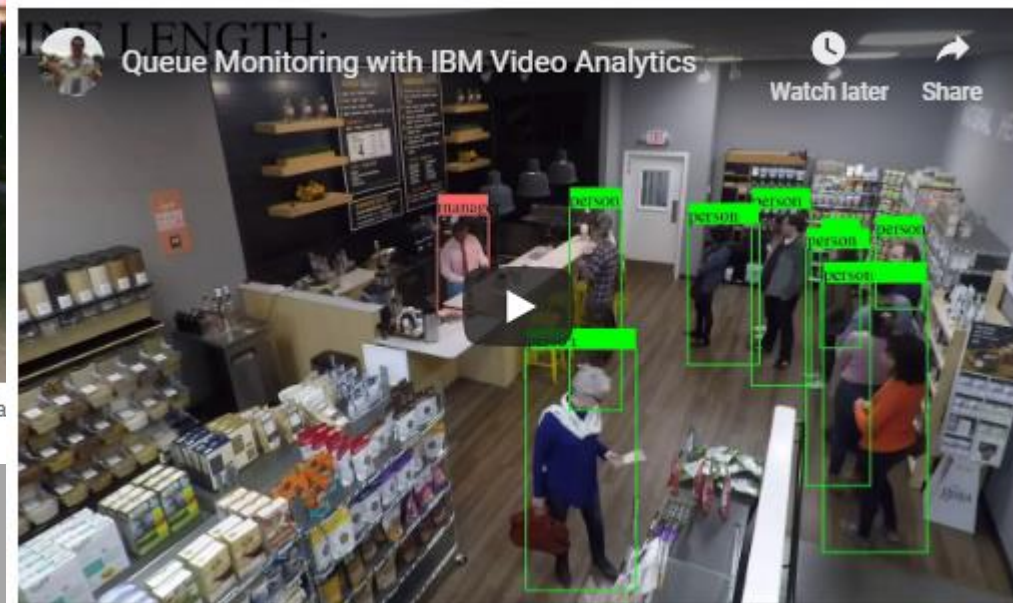
## *Intelligent 'Robot Surveillance' Poses Threats, A.C.L.U. Study Says*

Advances in artificial intelligence could supercharge surveillance cameras, allowing footage to be constantly monitored and instantly analyzed.

New York Times  
14 June 2019



A surveillance camera operated by the New Orleans Police Department fixed to a Bourbon Street. William Widmer for The New York Times



# Example Applications of Neural Networks

The neural network removes human bias (and insight)

IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011 150  
ISSN (Online): 1694-0814  
www.IJCSI.org

## Artificial Neural Networks in Medical Diagnosis

Qeethara Kadhim Al-Shayea

### The use of Neural Networks for the estimation of oceanic constituents based on the MERIS instrument

Daniel Buckton, Eon O'Mongain & Sean Danaher  
Pages 1841-1851 | Published online: 25 Nov 2010

Download citation <http://dx.doi.org/10.1080/014311699212515>

noise and variability models. The advantage of NNs is that they not only achieve higher retrieval accuracy than more traditional techniques such as band ratio algorithms, but they also allow the inclusion of usually superfluous or unused information, such as geometric parameters and atmospheric visibility.

# Training the Neural Network (1)

*The essence of a neural network is that it can “learn” from available data.* This is called *training* the NN. The NN has to “*learn*” what weighting functions will generate the desired output from the input.

Training can be done by *backpropagation of errors* when known inputs are compared with known outputs. We feed the NN various inputs along with the correct outputs, and let the NN objectively adjust its weights until it can reproduce the desired outputs.

The Java applet from [www.qub.ac.uk/mgt/intsys/perceptr.html](http://www.qub.ac.uk/mgt/intsys/perceptr.html) illustrates how a simple NN is trained by backpropagation.





# Things to Note

The NN was able to 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.

## Training the Neural Network (2)

Another way to train a NN is to view the NN as a complicated mathematical model that connects the inputs and outputs via equations whose coefficients (the weights) are unknown.

Then use a non-linear least squares fitting/search algorithm (e.g., Levenberg-Marquardt) to find the “best fit” set of weights for the given inputs and outputs (the training data).

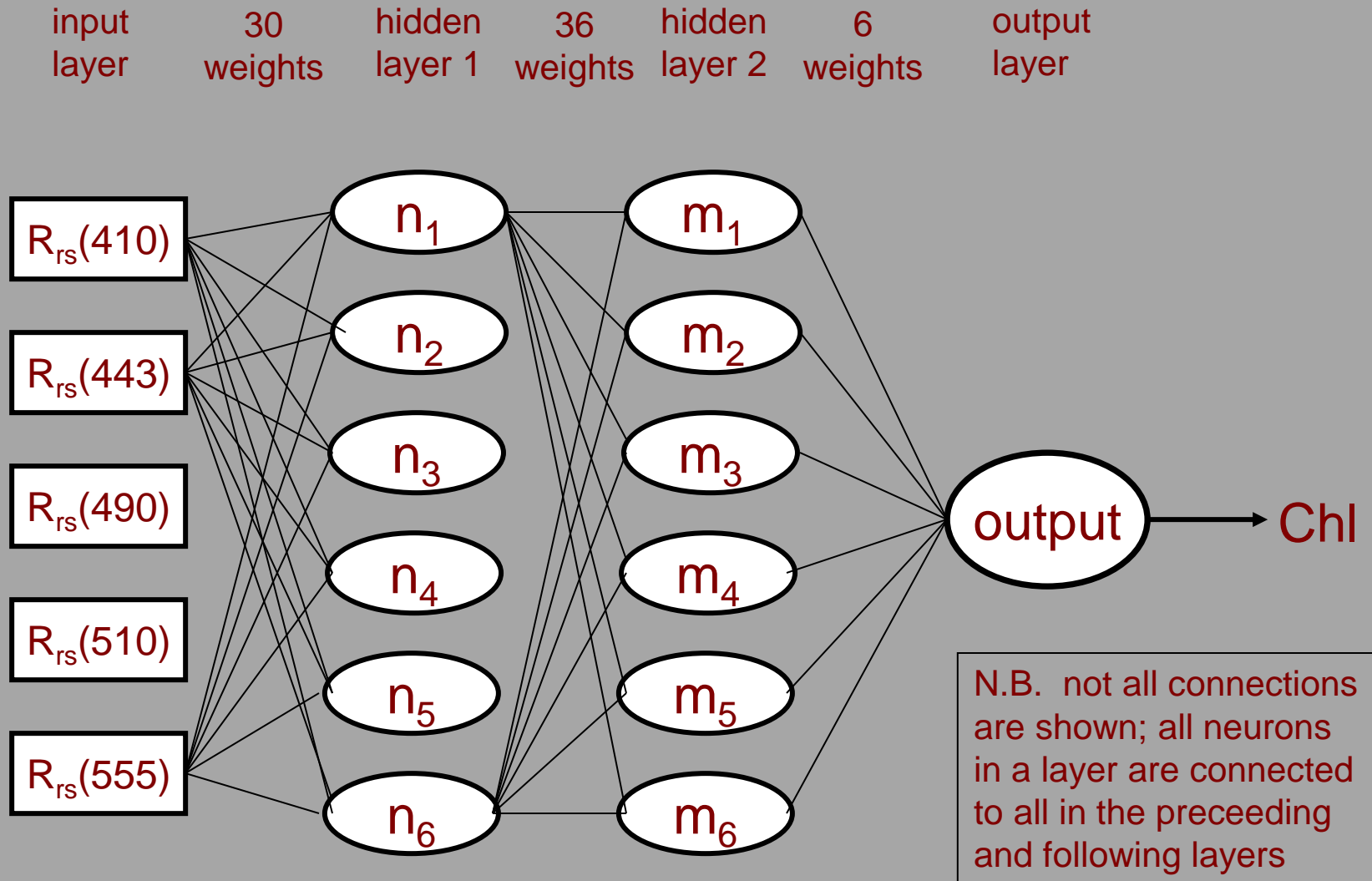
This makes it clear that NNs are just fancy regression models whose coefficients/weights are determined by fancy curve fitting to the available data (not a criticism!)

# An Example NN

From Ressom, H., R. L. Miller, P. Natarajan, and W. H. Slade, 1995. *Computational Intelligence and its Application in Remote Sensing*, in *Remote Sensing of Coastal Aquatic Environments*, R.L. Miller, C.E. Del Castillo, B.A. McKee, Eds.

- Assembled 1104 sets of corresponding  $R_{rs}$  spectra and  $Chl$  values from the SeaBAM, SeaBASS, and SIMBIOS databases.
- Constructed a NN with 5 inputs ( $R_{rs}$  at 5 wavelengths) and two hidden layers of 6 neurons each, and one output ( $Chl$ ).
- Partitioned the 1104 data points into 663 for training, 221 for validation, and 221 for testing the trained NN.
- The NN predictions of  $Chl$  using the testing data were compared with the corresponding  $Chl$  predictions made by the SeaWiFS OC4v4 band-ratio algorithm.

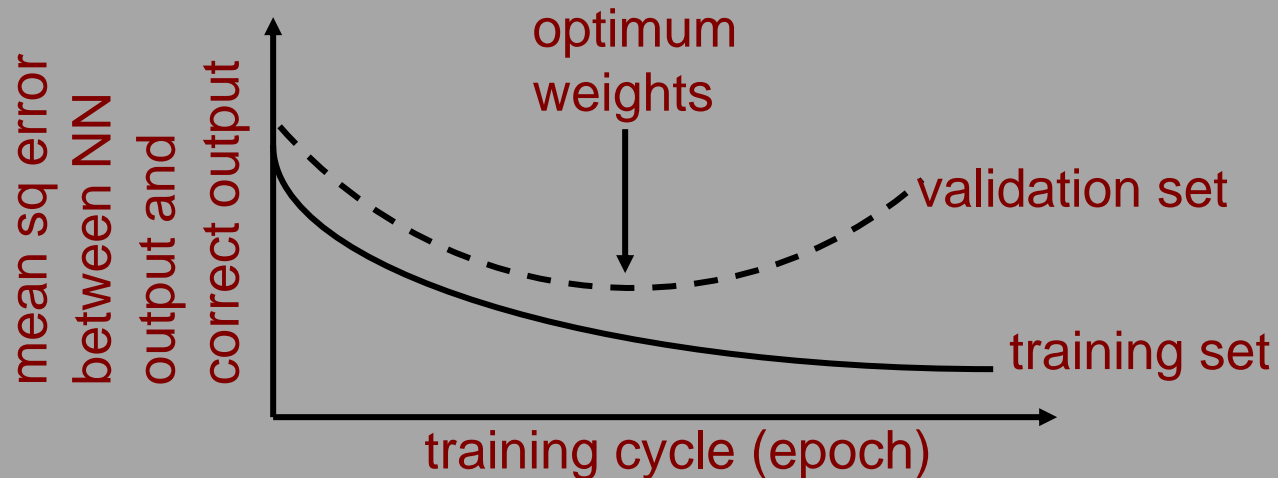
# The Resson et al. NN



# The Ressom et al. NN

Used two layers of 6 neurons, rather than one layer of 12, (for example), so that neurons can talk to each other (gives greater generality to the NN).

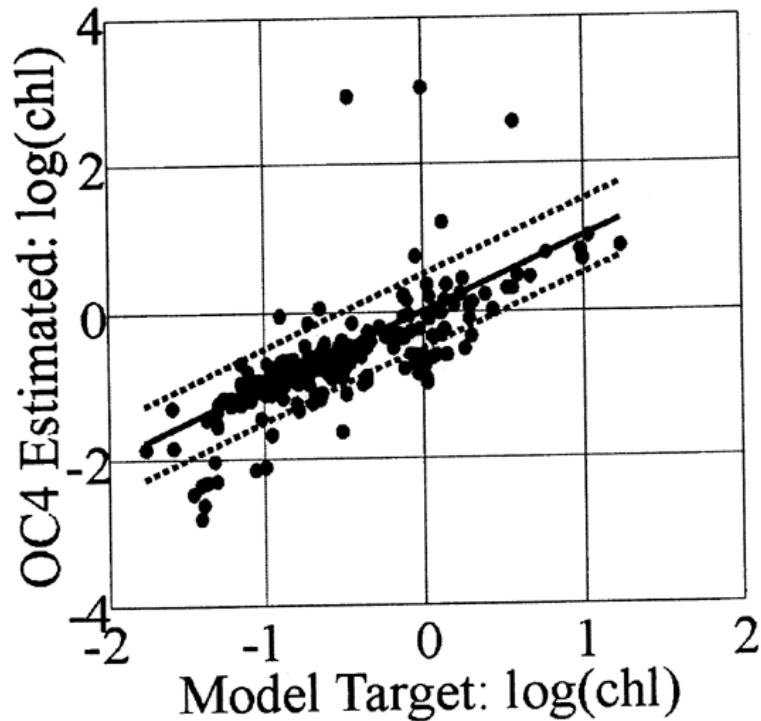
Training uses the training set for weigh adjustments, and the validation set to decide when to stop adjusting the weights.



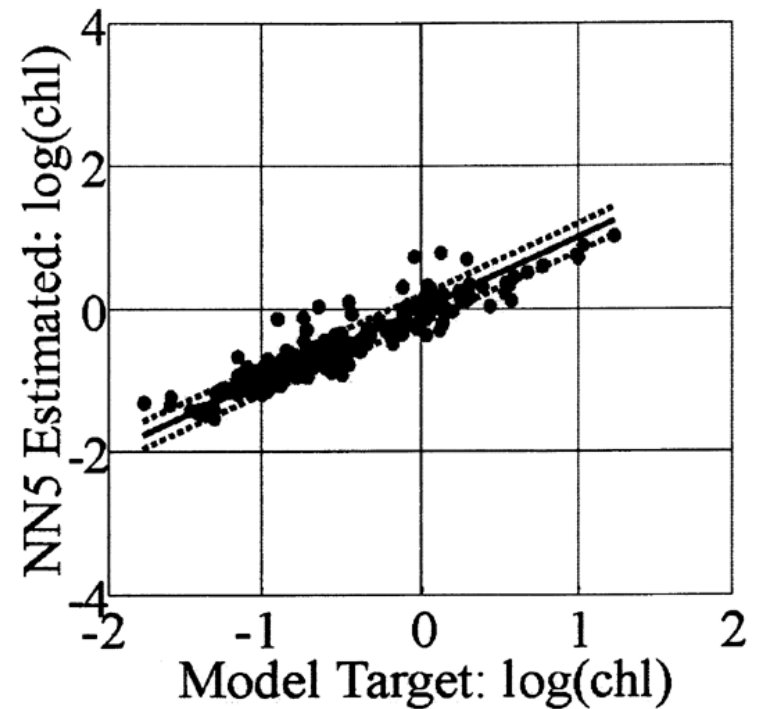
# NN vs. OC4v4 Performance

	Training Data (n=662)		Validation Data (n=221)		Testing Data (n=221)	
	$r^2$	RMS E	$r^2$	RMS E	$r^2$	RMS E
OC4	0.651	0.484	0.677	0.450	0.556	0.503
NN5	0.921	0.164	0.837	0.241	0.866	0.199

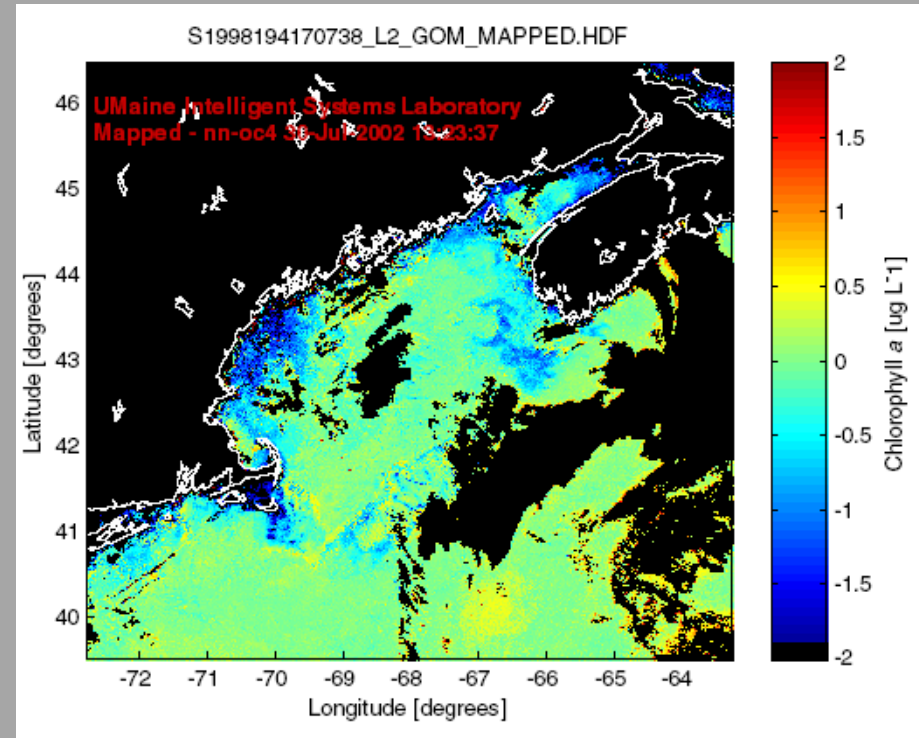
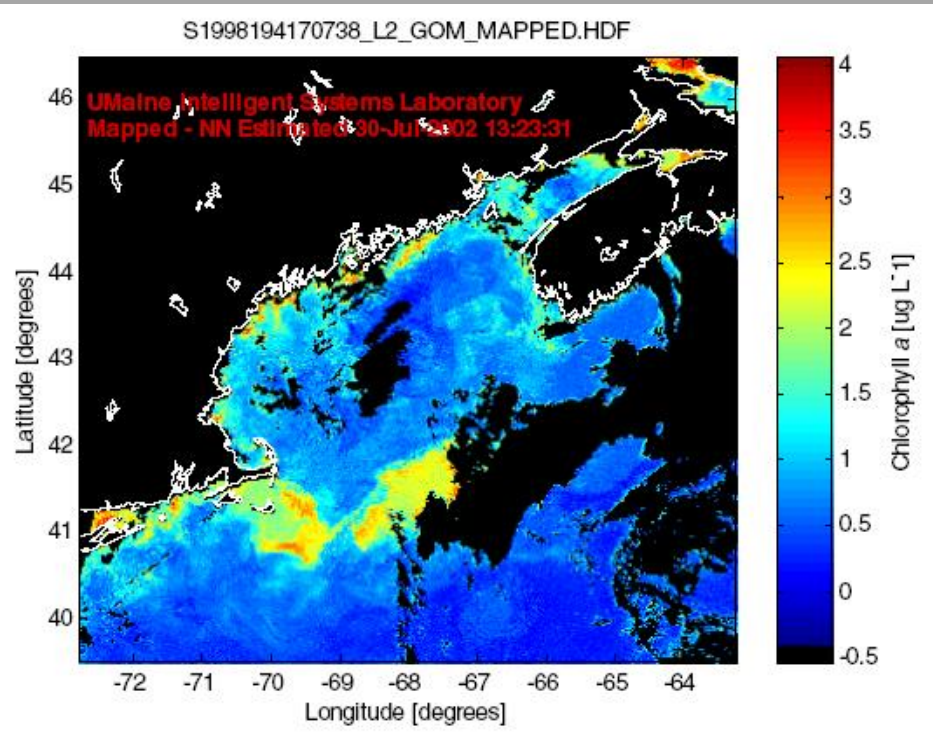
### OC4v4 Results



### NN5 Model Testing Results



# NN vs. OC4v4 Performance



Chl in the Gulf of Maine  
generated by applying a NN  
to SeaWiFS data

Difference in the NN and  
OC4 Chl values (NN-OC4)

from Slade, et al. Ocean Optics XVI



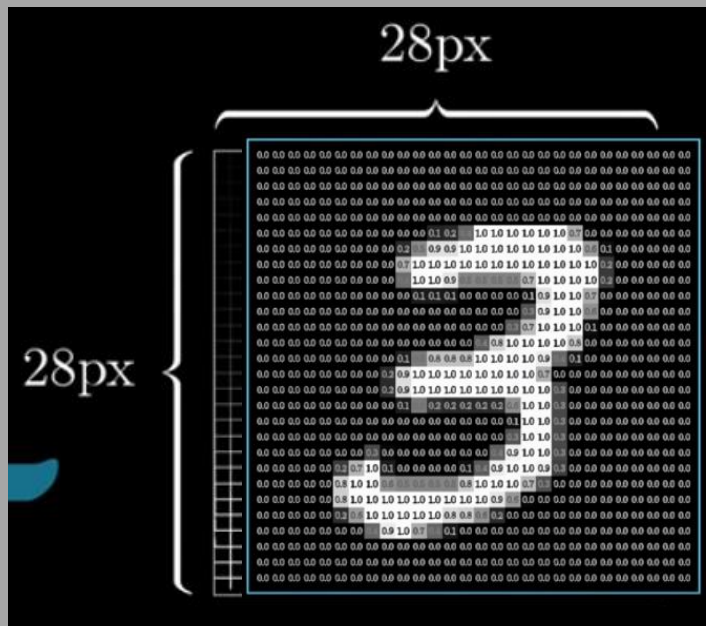
# For More on Neural Networks...



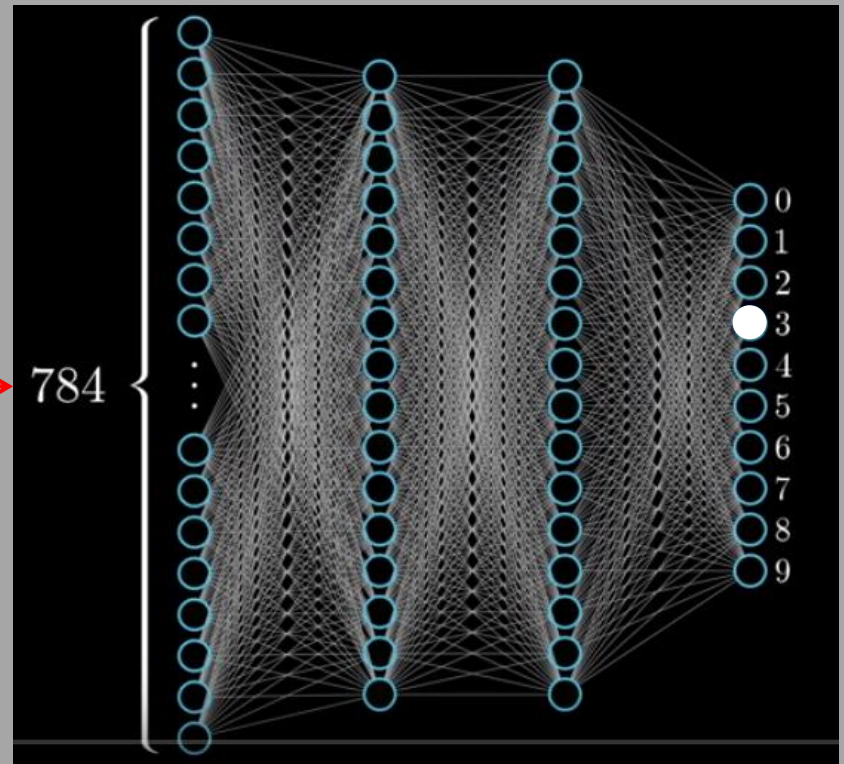
<https://www.3blue1brown.com/>

Youtube channel with excellent tutorials on many math different topics

Example shows how to use a NN to recognize handwritten numerals



28 x 28  
pixels =  
784  
neurons  
in the  
input  
layer



# Take-home Messages

Statistical methods for retrieving environmental information from remotely sensed data have been highly successful and are widely used, but...

- An empirical algorithm is only as good as the underlying data used to determine its parameters.
- This often ties the algorithm to a specific time and place. An algorithm tuned with data from the North Atlantic probably won't work well in Antarctic waters because of differences in the phytoplankton, and an algorithm that works for the Yellow Sea in summer may not work there in winter.
- The statistical nature of the algorithms often obscures the underlying biology or physics.

# Take-home Messages

Band-ratio algorithms remain operationally useful, but they have been milked for about all they are worth (IMHO). Note that band ratio algorithms throw away magnitude information in the  $R_{rs}$  spectra, and they may not use information at all available wavelengths.

New statistical techniques such as neural networks are proving to be very powerful, as are other techniques such as spectrum matching and semi-analytical techniques.



Limestone (early-mid Cambrian, 505-525 Myr old) boulder with fossil algal mats, Grand Canyon, photo by Curt Mobley

# Lava Falls, Grand Canyon



# Lava Falls, Grand Canyon



# Lava Falls, Grand Canyon



# Lava Falls, Grand Canyon





# Lava Falls, Grand Canyon

