OCEAN COLOUR AND BIOGEOCHEMICAL/ECOSYSTEM MODELS

Stephanie Dutkiewicz
Massachusetts Institute of Technology
MODIS TRUE COLOUR IMAGE 02/02/2011: Tropical cyclone Yasi 2011

What is happening to Great Barrier Reef?

Baird et al, ESM, 2015
% time that there is missing data in GlobColor (2002-2006)

What is impact of this missing data? e.g. on calculating spring bloom metrics?

Coles et al, JGR, 2012
Trend in SeaWiFs Chl 1997-2007

Is this “trend” a signature of anthropogenic driven climate change?
What causes differences in quantum yield?

What is the impact of missing data?

How to maintain monitoring under clouds?

Is this a climate change trend?

What drives the interannual variability?

What if clouds cover your region of interest?

Uncertainty from limited data to construct OC4 Chl algorithm?
NUMERICAL MODELS CAN BE USEFUL

What causes differences in quantum yield?

What is impact of missing data?

How to maintain monitoring under clouds?

Is this a climate change trend?

What drives the interannual variability?

What if clouds cover your region of interest?

Uncertainty from limited data to construct OC4 Chl algorithm?
Bachelor of Science (Physics) – University of Miami, FL, USA

PhD (Physical Oceanography) – University of Rhode Island, RI, USA

Post-doc – MIT, USA
  ecosystem model, SeaWiFs output

  carbon cycle, coupled climate modelling, iron cycle, adjoint modeling

  more complex ecosystem modeling, ecological theory –
  biodiversity, biogeographical provinces

Improving treatment of light/pigments in ecosystem model ....
OUTLINE

- Introduction to forward numerical biogeochemical /ecosystem models

- Why/How to use model output
  (OC example studies)

- Introduction to data assimilation
  (OC example studies)

- Capturing phytoplankton diversity
OUTLINE

- Introduction to forward numerical biogeochemical/ecosystem models
- Why/How to use model output (OC example studies)
- Introduction to data assimilation (OC example studies)
- Capturing phytoplankton diversity
Model: “simplified representation of real system or process”
Model: “simplified representation of real system or process”

often described by a set of equations
Phytoplankton

Zooplankton

Fish

Sharks/mammals/birds

Detritus

Benthic organisms

Corals

Bacteria

Chl input

Nutrients

Sunlight

Stored in deep ocean
sunlight

phytoplankton

zooplankton

nutrients

bacteria

detritus

corals

benthic organisms

sharks/mammals/birds

fish

big fish

stored in deep ocean
<table>
<thead>
<tr>
<th>Scale</th>
<th>Trophic Level</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>cell level</td>
<td><strong>Lower trophic</strong></td>
<td>0-D</td>
</tr>
<tr>
<td>individual based</td>
<td><strong>Higher trophic</strong></td>
<td>1-D</td>
</tr>
<tr>
<td>continuum</td>
<td><strong>End-to-end</strong></td>
<td>2-D</td>
</tr>
</tbody>
</table>

**Dimension:** 3-D (regional/global)

**Time-stepping**
**MODELLING PHYTOPLANKTON**

**Biomass of Phytoplankton (B):**

\[
B_{\text{at time 2}} = B_{\text{at time 1}} + (\text{growth} - \text{losses} + \text{transport}) \times \text{timestep}
\]

- **growth**: nutrient limitation, light, temperature, ...
- **losses**: grazing, viruses, respiration, cell death, sinking, ...
- **transport**: advection by currents, vertical mixing, ...
rate of change of biomass = growth - loss - transport

\[
\frac{dB}{dt} = \mu B - gZ - mB - \nabla \cdot \vec{u} B + \nabla \cdot (\hat{K} \nabla B)
\]
rate of change of biomass = growth - loss - transport

\[
\frac{dB}{dt} = \mu B - gZ - mB - \nabla \cdot \vec{u}B + \nabla \cdot (\vec{K}\nabla B)
\]

functional forms for each biological term are “approximations”; there is considerable research in how to parameterize each
RESOLUTION:
Horizontal: Coarse: > 1 degree
    Fine: 10’s km

Vertical: usually finer in surface (m’s, and
    coarser 100m at depth)

(also different types of grids – isopynal etc)

Grid size: “resolution” depends on model purpose
MODELLING PHYTOPLANKTON

**Computer code**

**Equations**

\[
\frac{dR}{dt} = -\mu \frac{R}{R + \kappa_R} B + rD - \nabla \cdot \tilde{u} R + \nabla \cdot (\tilde{K} \nabla R)
\]

\[
\frac{dB}{dt} = \mu \frac{R}{R + \kappa_R} B - g \frac{B}{B + k_B} Z - mB - \nabla \cdot \tilde{u} B + \nabla \cdot (\tilde{K} \nabla B)
\]

\[
\frac{dZ}{dt} = \gamma g \frac{B}{B + k_B} Z - mZ - c Z^2 - \nabla \cdot \tilde{u} Z + \nabla \cdot (\tilde{K} \nabla Z)
\]

\[
\frac{dD}{dt} = S(D) - rD - \nabla \cdot \tilde{u} D + \nabla \cdot (\tilde{K} \nabla D)
\]
MODELLING PHYTOPLANKTON

Physics: e.g. velocity, mixing, temperature (e.g. MITgcm, ROMS, MOM etc)

Biogeochemistry: e.g. Carbon, nutrients, DOM, POM

Ecosystem: e.g. Phytoplankton (C, Chl), zooplankton

Movie credit: Oliver Jahn and Mick Follows
• Modeling phytoplankton
  - equations, parameterization
  - grid, computer code, computer simulations

• Biogeochemical/Ecosystem models include:
  - physics
  - biogeochemistry
  - ecosystem
SOME TERMINOLOGY (1)

- Physical model – temperature, velocities, mixing
- Biogeochemical model – cycling of elements such as C,N,P
- Ecosystem model – life and death of organisms

- State Variable – things (“concentrations”) that the model steps forward in time (e.g. phytoplankton biomass)

- Grid resolution – spatial size of grids (horizontal/vertical)

- Initial conditions – state the model starts from (usually climatologies)
  - Forcing fields – boundary conditions, incoming conditions that are assumed as knowns

- Hindcast model – simulates last few decades
- Climate model – usually has hindcast and future components (CMIP5 / CMIP6)
EVALUATION OF MODEL WITH OCEAN COLOUR

Model’s evaluated output against ocean colour products – particularly Chl (Level 3, gridded), but also primary production, sometimes Kd, PFTs

**GFDL’s ESM2:** Dunne et al, J Clim, 2011

**MIROC-ESM:** Watanabe et al, GMD, 2011

**BOGM, Doney et al, JMS, 2008**
- Introduction to forward numerical biogeochemical /ecosystem models

- Why/How to use model output (OC example studies)

- Introduction to data assimilation (OC example studies)

- Capturing phytoplankton diversity
WHY USE MODEL OUTPUT

Surface Chl (mg/m3)

Movie credit: Oliver Jahn
OCEAN ACIDIFICATION

WHY USE MODEL OUTPUT

Chl at 50m (mg/m³)

Movie credit: Oliver Jahn
Just like any other “dataset”

- gridded and no missing points
- but with caveats
“... all models are wrong, but some are useful”

George Box
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

% time that there is missing data in GlobColor (2002-2006)

What is impact of this missing data?
e.g. on calculating spring bloom metrics?

Coles et al, JGR, 2012
Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics

Harriet Cole,¹ Stephanie Henson,² Adrian Martin,² and Andrew Yool²

Observations: GlobColor (SeaWiFS, MERIS, MODIS synthesis)

Model: NOBM
Observations: GlobColor (SeaWiFS, MERIS, MODIS synthesis)

Model: NOBM

Spring Bloom Metrics:

Peak (maximum)
Initiation (5% above annual median)
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

Why are timing metric different when calculated with gappy data?

[Graph showing Chl (mg/m³) over months]
Why are timing metric different when calculated with gappy data?

**Peak:** will be when highest Chl is captured, not necessarily highest actual
Uncertainty (in days) of bloom peak

white areas: no discernable seasonal cycle

Uncertainty (in days) of bloom initiation much worse

Coles et al, JGR, 2012
What does this study indicate?
- Will be tough to understand interannual variability of timing spring bloom (and accompanying recruitment, export etc issues)
- Trends in phenology could be missed (within uncertainty estimates)
Take home:
- Model can help understand uncertainty in ocean colour timeseries studies (i.e. not just spring bloom) from lack of data
Satellite-detected fluorescence reveals global physiology of ocean phytoplankton

M. J. Behrenfeld¹, T. K. Westberry¹, E. S. Boss², R. T. O’Malley¹, D. A. Siegel³, J. D. Wiggert⁴, B. A. Franz⁵, C. R. McClain⁶, G. C. Feldman⁵, S. C. Doney⁷, J. K. Moore⁷, G. Dall’Olmo¹, A. J. Milligan¹, I. Lima⁸, and N. Mahowald⁸

Chlorophyll fluorescence

\[ F = Chl \times \langle a^*_\text{ph} \rangle \times iP iRA \times \phi \]

Estimate quantum yield \( \phi \)
Ocean Colour Derived Quantum Yield

Suggests difference in physiological state:
But what causes this?
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

BOGM model (Moore et al 2006):
Global, coarse resolution

\[
\mu = \mu_{\text{max}} f(T, I) \min\left( \frac{NO_3}{NO_3 + \kappa_{\text{no}_3}}, \frac{PO_4}{PO_4 + \kappa_{\text{po}_4}}, \frac{Fe}{Fe + \kappa_{\text{fe}}} \right)
\]

Behrenfeld et al., BG, 2009
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

Higher resolution Indian Ocean model

(Wiggert et al 2006)
Take home:
- By having more mechanistic details (here growth limitation) models can help understand ocean colour results
- Models do not work as well in all locations – using multiple models can be useful.
Physical drivers of interannual chlorophyll variability in the eastern subtropical North Atlantic

M. V. Pastor, J. B. Palter, J. L. Pelegrí, and J. P. Dunne

What drives the interannual variability?
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

Pastor et al, JGR, 2013

TOPAZ model
Take home:
- Interannual variability embedded in multidecadal oscillation (need the model simulation to extend the “observations”)
- Variability is driven by alterations in nutrient supply which in turn is driven by changes in upwelling water (processes that can not be detected using satellite observations)
Potential for models to inform ocean colour community:

• Uncover processes driving observations
• Explore uncertainties in observations (gaps in data etc)
• Extend dataset in space and time
• Place variability in large context

Debunk Model Myths:

• *You do not need to understand all equations and output*
• *Output is (or can be made) available*
• *Models are inherently wrong – but they can be useful*
OUTLINE

- Introduction to forward numerical biogeochemical /ecosystem models
- Why/How to use model output (OC example studies)
- Introduction to data assimilation (OC example studies)
- Capturing phytoplankton diversity
DATA ASSIMILATION

- Combine observation with model
- Bring model into consistency with observations
- “Best” of model and observations
Some different types of data assimilation:

- Sequential methods
  (e.g. Kalman Filter, Ensemble Kalman Filter)

- Variational methods
  (e.g. Adjoint, 4D Var)
DATA ASSIMILATION

A diagram showing the relationship between state and time with data points at times $t_1$, $t_2$, and $t_3$. The data points are connected to an arrow labeled "data."
DATA ASSIMILATION

A diagram illustrating the concept of data assimilation in the context of ocean acidification. The graph shows the state over time, with data points at specific time intervals ($t_1$, $t_2$, $t_3$) and an unconstrained model line connecting these points. The data points are indicated by red circles with error bars, and the model line is depicted with a blue line.
Assimilation: Sequential Method (e.g. Kalman Filter)
DATA ASSIMILATION: EXAMPLE

- 4 fold decrease in bias;
- Improves non-assimilated processes such as primary production
- Does not necessarily conserve mass; can also lead to misfits between nutrients and Chl (though see Rousseaux et al 2012)
DATA ASSIMILATION

- For providing state estimates/re-analyses
- For parameter optimization
Some different types of data assimilation:

- **Sequential methods**
  (e.g. Kalman Filter, Ensemble Kalman Filter)

- **Inverse methods**
  (e.g. Adjoint, 4D Var)
Data Assimilation

Least square fit of model to observations:

\[ J = \sum_i W_i^2 (\text{model}_i - \text{data}_i)^2 \]

- \( J \) - cost function
- \( W_i \) - weighting function (error estimates)
State Estimation: Adjoint Method

- essentially the “backward” model
- efficiently computes sensitivity of model (cost function) to perturbations in parameters/initial conditions/forcing fields

\[ J = \sum_i W_i^2 (\text{model}_i - \text{data}_i)^2 \]

Use adjoint to find gradient \( \frac{dJ}{dx} \) to iteratively minimize \( J \)

\( x \) can be:
- initial conditions
- forcing fields
- model parameters
DATA ASSIMILATION

Assimilation: Variational Method (e.g. Adjoint)
The assimilation of satellite-derived data into a one-dimensional lower trophic level marine ecosystem model

Yongjin Xiao¹ and Marjorie A. M. Friedrichs¹
Forward Model

Adjoint Model after parameter estimation to be consistent with ocean colour data

Xiao and Friedrichs, JGR, 2016
SOME TERMINOLOGY (3)

- Forward model – not data assimilation
- Data Assimilation – produces “analysis”, “re-analysis” and/or parameter estimations
- Operational modelling – produce “products” for real-time or short term uses
- Inverse modeling – know the results, figure out parameters/initial conditions/forcing fields to get there (e.g. adjoint)
- Cost function – least square difference between model and data (goal of assimilation is to reduce cost function)
- Forecast: after re-initializing for “today” run forward (days) for prediction
- Reanalysis/State Estimation - data assimilated fields (e.g. Chl)
Assimilation: Sequential Method
(e.g. Kalman Filter)
Data Assimilation

Assimilation: Sequential Method
(e.g. Kalman Filter)
DATA ASSIMILATION

Assimilation: Sequential Method (e.g. Kalman Filter)
Assimilation: Sequential Method
SUMMARY (3)

- Types of data assimilation:
  - sequential methods
  - inverse methods (e.g. adjoint)

- Provide analysis/re-analysis/forecast: but know limitations

- Can also provide parameter estimates

- Physical fields from data assimilation:
  ECCO (adjoint)
  OTHERS.....

- Physical and biogeochemical/ecosystem assimilation eventually
And what sort is more appropriate for my question/interest?
HOW TO GET MODEL OUTPUT: EXAMPLES

NASA MODEL:

NOBM (Gregg, JMS, 2008)
Assimilates Chl,
Hindcast available

Data available from Giovanni
Atlantic - European North West Shelf - Ocean Biogeochemistry Analysis and Forecast

The Forecasting Ocean Assimilation Model 7km Atlantic Margin model (FOAM AMM7) is a coupled hydrodynamic-ecosystem model, nested in a series of one-way nests to the Met Office global ocean model. The hydrodynamics are supplied by the Nucleus for European Modelling of the Ocean (NEMO) with the 3DVar NEMOVAR system used for the assimilation of sea surface temperature data. This is coupled to the European Regional Seas Ecosystem Model (ERSEM), developed at Plymouth Marine Laboratory (PML). ERSEM based models have been used operationally to forecast biogeochemistry in the region for a number of years (Siddorn et al., 2006).

Atlantic- European North West Shelf- Ocean Biogeochemistry Reanalysis from METOFFICE (1985-2012)

The reanalysis covers the period January 1985 until July 2012 and is based upon the Forecasting Ocean Assimilation Model 7km Atlantic Margin Model (FOAM AMM7). This is a coupled hydrodynamic-ecosystem model of the North West European shelf forced at the surface by ERA-interim winds, atmospheric temperature, and precipitation fluxes. Horizontal boundary conditions were provided by the NOC global reanalysis prior to 1989 and by the GloSea reanalysis thereafter. Boundary conditions in the Baltic Sea came from the IOM-GETM model. E-Hype data were used for river inputs. Hydrodynamic calculations were performed by the Nucleus for European Modelling of the Ocean (NEMO) system, while the 3DVar NEMOVAR system was used for the assimilation of sea surface temperature data. This is coupled to the European Regional Seas Ecosystem Model (ERSEM), developed at Plymouth Marine Laboratory (PML) (Blackford et al., 2004). ERSEM based models have been used operationally to forecast biogeochemistry in the region for a number of years (Siddorn et al., 2007). Nutrient boundary conditions are provided by World Ocean Atlas climatology (Garcia et al., 2010) and river nutrient and sediment concentrations from the operational climatology. Outputs are provided both as monthly means and as daily 25 hour, de-tided, averages.

WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

What happens if you want to go fishing today here?
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

Pastor et al, JGR, 2013

TOPAZ model
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

BOGM model (Moore et al 2006):

\[ \mu = \mu_{\text{max}} f(T, I) \min \left( \frac{NO_3}{NO_3 + \kappa_{\text{no3}}}, \frac{PO_4}{PO_4 + \kappa_{\text{po4}}}, \frac{Fe}{Fe + \kappa_{\text{fe}}} \right) \]

Behrenfeld et al., BG, 2009
WHY/HOW TO USE MODEL OUTPUT: EXAMPLES

Uncertainty (in days) of bloom peak

white areas: no discernable seasonal cycle

Uncertainty (in days) of bloom initiation much worse

Coles et al, JGR, 2012
Example of non-assimilated models:

BOGM (Moore et al., 2007) -- NCAR

PISCES (Aumont et al, GMD, 2015)

Darwin (Dutkiewicz et al, BG, 2015)

Topaz (Dunne et al, J Clim, 2011) – GFDL

(Kearney et al, Ecol Model., 2012) – GFDL (end to end foodweb model)

…. many others…..
- Introduction to forward numerical biogeochemical /ecosystem models
- Why/How to use model output (OC example studies)
- Introduction to data assimilation (OC example studies)
- Capturing phytoplankton diversity
MODELING DIVERSITY OF PHYTOPLANKTON

sunlight

phytoplankton

zooplankton

fish

sharks/mammals/birds

big fish

stored in deep ocean

detritus

nutrients

corals

benthic organisms

stored in deep ocean
Phytoplankton diversity matters:

- As base of foodweb
- For carbon cycling
- For resilience of ecosystem
Phytoplankton differ in terms of:

- size
- nutrient uptake
- biogeochemical function
- temperature tolerance
- pigment
- predation avoidance
- morphology
- chain/colony formation
- symbiosis
Phytoplankton differ in terms of:

- size
- nutrient uptake
- biogeochemical function
- temperature tolerance
- pigment
- predation avoidance
- morphology
- chain/colony formation
- symbiosis
MODELING DIVERSITY: NUTRIENT UPTAKE

\[ \mu_{\text{max}} \frac{R}{R + \kappa_R} \]

Cermeno et al, MEPS, 2011

![Graph showing nutrient uptake and growth rate](image)
MODELING DIVERSITY: NUTRIENT UPTAKE

\[ \mu_{\text{max}} \frac{R}{R + \kappa_R} \]

Cermeno et al, MEPS, 2011
MODELING DIVERSITY: NUTRIENT UPTAKE

- K-strategist (gleaner)
- dominates in stable, low nutrient conditions (resource control theory)

- r-strategist (opportunist)
- dominates in high nutrient conditions (fastest net growth important)

Biomass (mgC/m3)

Dutkiewicz et al, GBC, 2005
MODELING DIVERSITY: NUTRIENT UPTAKE

Model Examples:
Chai et al, 2002
Moore et al, 2002
Aumont et al, 2005
Dutkiewicz et al, 2005; 2009

most CMIP5 models
Phytoplankton functional types (PFT)

- calcifiers
- silicifiers
- N$_2$ fixers
- mixotrophs

Le Quere et al, Glob Change Bio, 2005
MODELING DIVERSITY: BIOGEOCHEMICAL FUNCTION

Phytoplankton functional types (PFT)

Model Examples:
- Gregg and Casey, DSR, 2007
- Le Quere et al, 2005
- Moore et al, 2004
- Hood et al, 2004
- Dutkiewicz et al, 2015
MODELING DIVERSITY: SIZE

Finkel et al, JPR, 2009
MODELING DIVERSITY: SIZE

Maximum growth rate

\[ a V^b \]

Nutrient half saturation

Model Examples:
- Moloney and Fields, 1991
- Armstrong, 1999
- Baird and Suthers, 2007
- Banas et al, 2011
MODELING DIVERSITY: SIZE

Satellite derived
(Hirata et al, 2011)

model

Phytoplankton differ in terms of:

- size
- nutrient uptake
- biogeochemical function
- temperature tolerance
- pigment
- predation avoidance
- morphology
- chain/colony formation
- symbiosis

Clayton et al, L+OFE, 2014
Movie credit: Oliver Jahn
Phytoplankton diversity matters for foodweb, carbon cycle

Most models include between 2 and 6 “phytoplankton functional types” (PFTs) – thus link to some OC estimates (but may not be the same choices)

Some models model range of sizes using allometric scaling with volume of various traits – thus link to other OC estimates (e.g. abundance based)

But there are sizes within functional groups (e.g. diatoms), and many other reasons for phytoplankton diversity
What causes differences in quantum yield?

What is impact of missing data?

How to maintain monitoring under clouds?

What drives the interannual variability?

What if clouds cover your region of interest?

Is this a climate change trend?

Uncertainty from limited data to construct OC4 Chl algorithm?

NUMERICAL MODELS CAN BE USEFUL