# Building Custom Ecosystem Models with PALEOtoolkit

Sara Sjosten, Stuart Daines

# Schedule

- **Introductions:** workshop leaders, participants, goals and activities
- **Introduction to modelling:** theory, examples, evaluation process
- **Group exercise:** what processes in my work could be modelled?
- 9.30-9.45am: Break
- **Introduction to the PALEOtoolkit:** motivation and approach
- **Tutorial 1:** group walk-through of NP (primary production) model
- **Tutorial 2:** hands-on exercise creating NPZ
- **Closing:** reflections, next steps

# Workshop Goals

Participants will leave with...

- an understanding of how to use models to **generate knowledge**;
- a firmer sense of **what modelling is**, if new to modelling;
- the ability to define models to **test hypotheses** about the Earth system;
- awareness of how **PALEOtoolkit** can be used to test Earth system models;
- confidence engaging with the **mathematics and code** used for models.

# What is modelling?

- **Cambridge dictionary:** Modelling is "the activity of using mathematical models (simple descriptions of a system or process) to do calculations or predict what might happen"

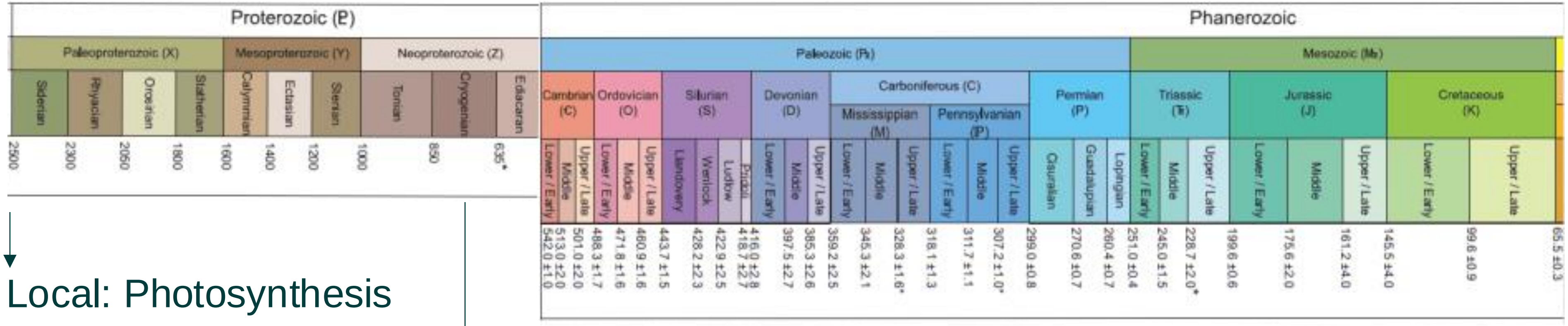
- **Dictionary.com:** A model is "a simplified representation of a system or phenomenon, as in the sciences or economics, with any hypotheses required to describe the system or explain the phenomenon, often mathematically"

# What is modelling for our purposes?

- **Scientific hypothesis** about a **process** or system of processes in the **Earth system** and how the system **evolves over time**
- **Mathematical expression** of the processes in terms of how it affects **physical matter**
- How the processes connect with Earth system **biogeochemical cycles** across different **temporal and spatial scales**
- Matter must be **conserved** across scales!
- Need to **predict some observable data** for testing!

# Intro to Modelling

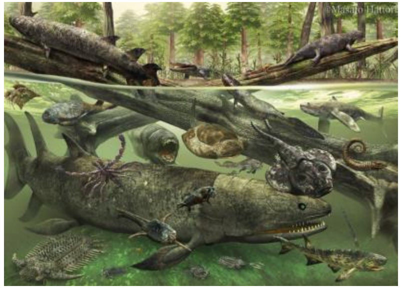
Example: Great Oxidation Event




Local: Photosynthesis
Global: GOE

Local: Sponge evolution
Global: Redox change

Local: Plant evolution
Global: LDME




Local: Biotic innovations
Global: Redox change

# Example: Great Oxidation Event

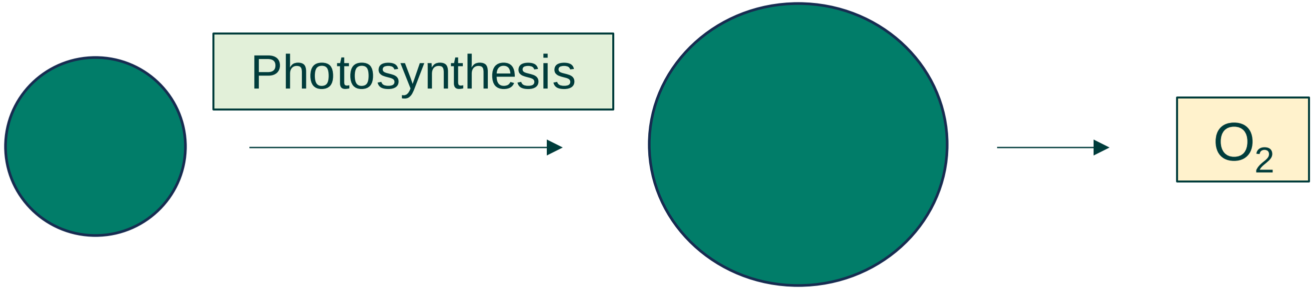
- Transition from the **methanogenic atmosphere** of early Earth to the current **oxygenic atmosphere** ~2.4 Gyr
- Transition driven by **evolution of oxygenic photosynthesis** ~2.7 Gyr
- Spatially and temporally **local biological process** in the ocean driving a permanent **system-level change** to the surface Earth

# Example: Great Oxidation Event

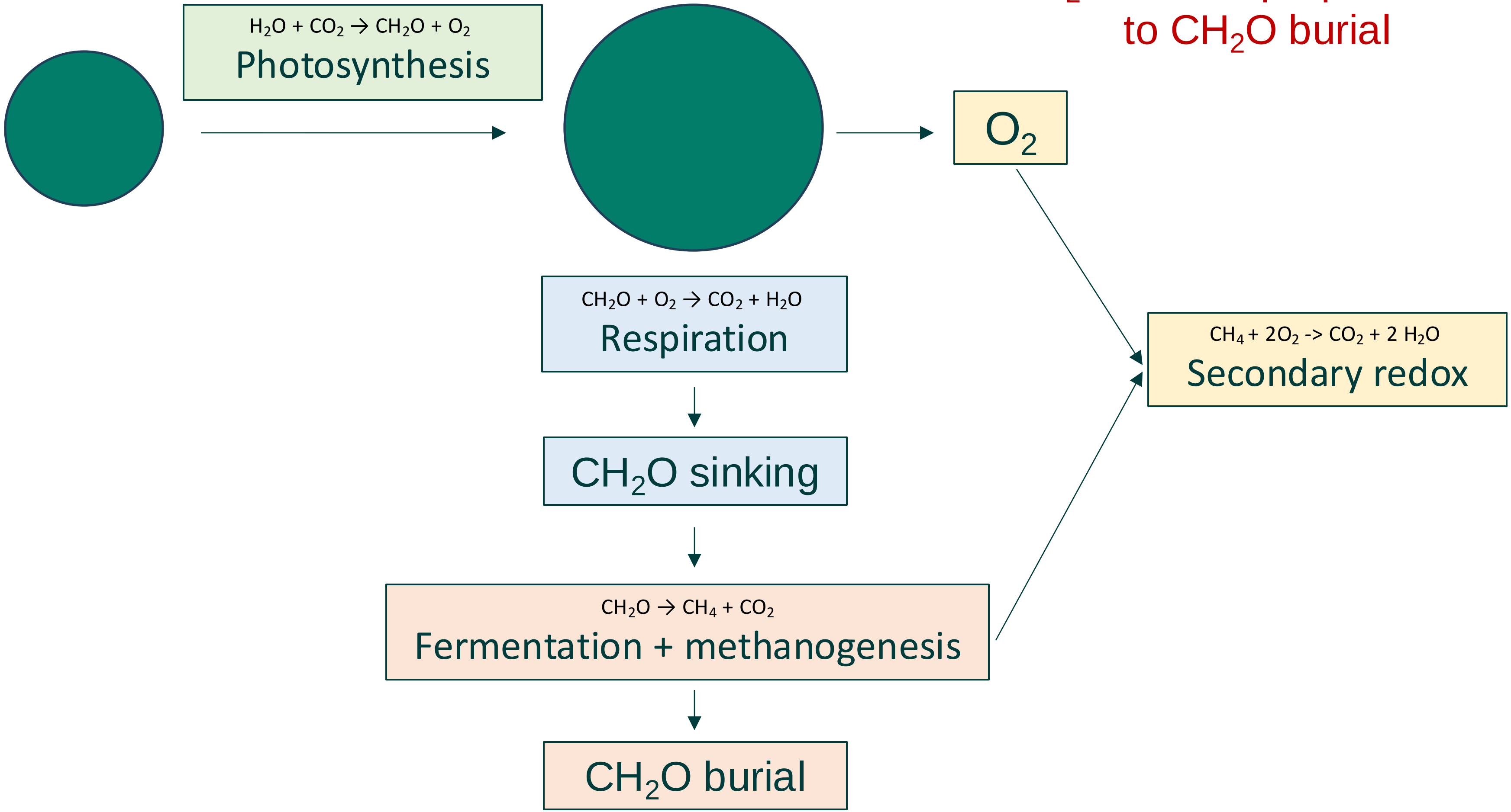
Photosynthesis $H_2O + CO_2 \rightarrow \text{“}CH_2O\text{”} + O_2$

Respiration $\text{“}CH_2O\text{”} + O_2 \rightarrow CO_2 + H_2O$

$O_2$ increase proportional to gain in biomass

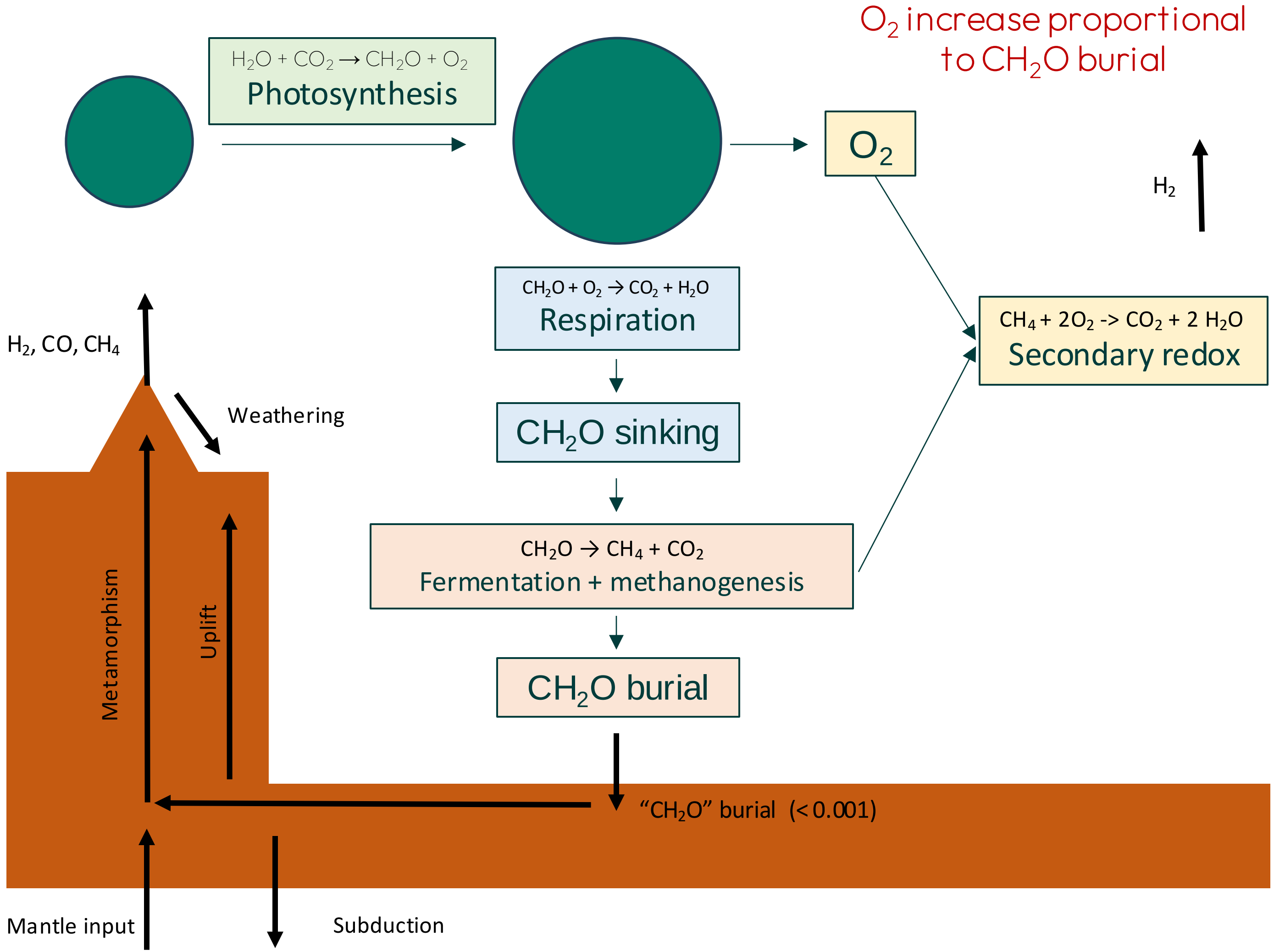


$O_2$ increase proportional to $CH_2O$ burial



If net $O_2$ < "volcanic" reduced gasses, atmos is anoxic
(redox budget closed by H escape)

If net $O_2$ > "volcanic" reduced gasses, atmos is oxic
(redox budget closed by oxidative weathering)



# Organic Matter Cycling

## Biology:

1. Photosynthesis: $CO_2 + H_2O \rightarrow CH_2O + O_2$
2. Aerobic respiration: $CH_2O + O_2 \rightarrow CO_2 + H_2O$
3. Methanogenesis: $2CH_2O \rightarrow CH_3COOH \rightarrow CH_4 + CO_2$
4. Methanotrophy: $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$

## Burial and tectonic recycling:

Only the organic material from (1) photosynthesis that escapes (2) aerobic respiration or (3) methanogenesis and (4) methanotrophy provides a net $O_2$ source:

1. Organic carbon burial: $CO_2 + H_2O \rightarrow CH_2O(\downarrow \text{burial}) + O_2$
2. Metamorphism and subduction: $CH_2O \rightarrow CH_4 + CO_2$ ($\uparrow$ "volcanic" gasses)
3. Oxidative weathering: $CH_2O(\uparrow \text{weathering}) + O_2 \rightarrow CO_2 + H_2O$

A+B+C is a null cycle (no net source or sink of $O_2$) if the amount of organics ($CH_2O$) in the crust is at steady state

# What Caused the Great Oxidation Event?

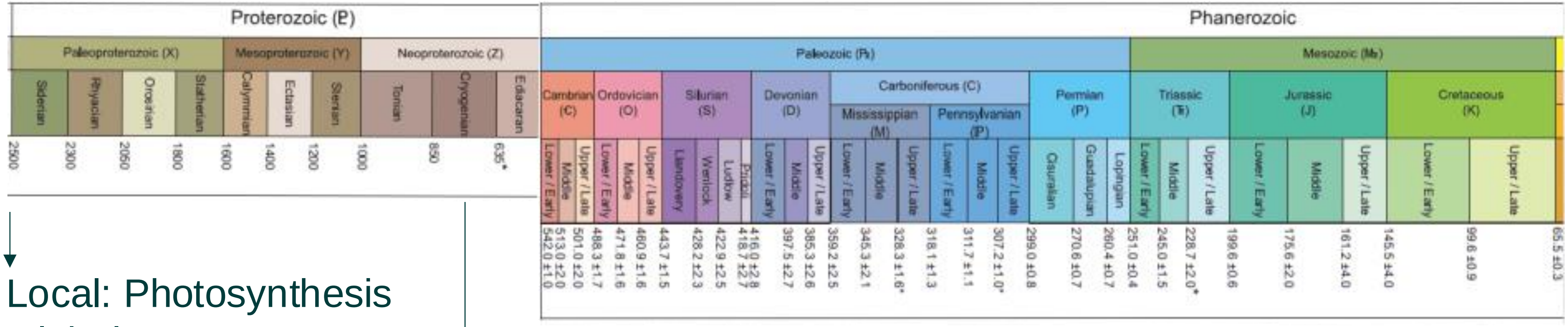
1. Increase in "$CH_2O$" burial? (oxygen source)
   a. Constrained by carbon isotope data ($\delta^{13}C$)
   b. 'Transient' (< 100My sediment cycle time)
2. Decrease in reductant input?
   a. Mantle degassing?
   b. Slow oxidation driven by $H_2$ escape?
   c. End of formation (and oxidation) of continents?

# What Processes to Consider?

Often useful to classify ecosystem processes into:

1. Autotrophic (biomass production)
   i. Photosynthesis, etc.
   ii. Often tractable, small molecule inorganic substrates
2. Heterotrophic (biomass degradation)
   i. Respiration, etc.
   ii. "$CH_2O$" -> $CH_4 + CO_2$ (disproportionation, without TEA)
   iii. Complicated, multiple organic carbon substrates
3. Secondary redox
   i. eg $CH_4 + 2O_2 \rightarrow CO_2 + 2\,H_2O$ (either in the biota or atmosphere)
4. Preservation and burial (net oxygen source)
   i. On long timescales
   ii. Poorly understood

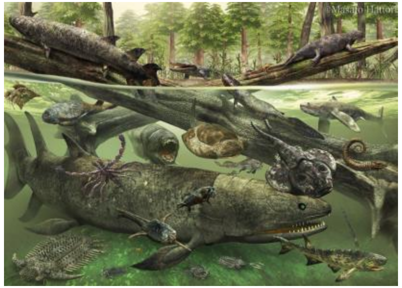
To construct a model we don't necessarily need to represent each part completely!




Local: Photosynthesis
Global: GOE

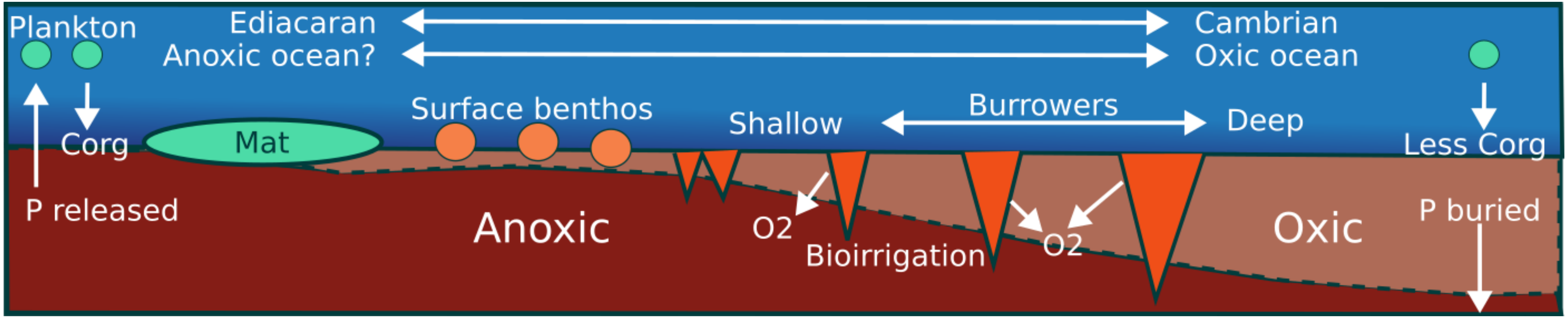
Local: Sponge evolution
Global: Redox change

Local: Plant evolution
Global: LDME

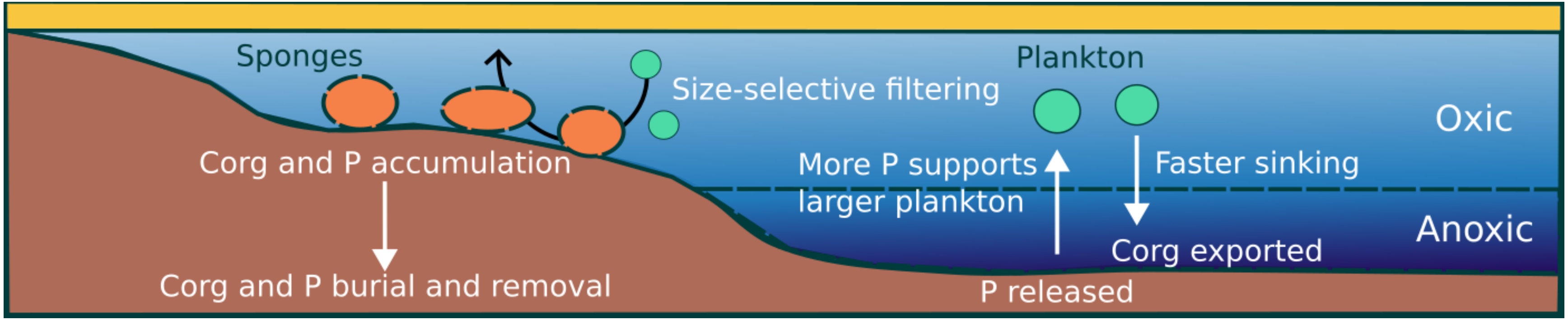


Local: Biotic innovations
Global: Redox change

# Cambrian Substrate Revolution




# Sponge size-selective filtering

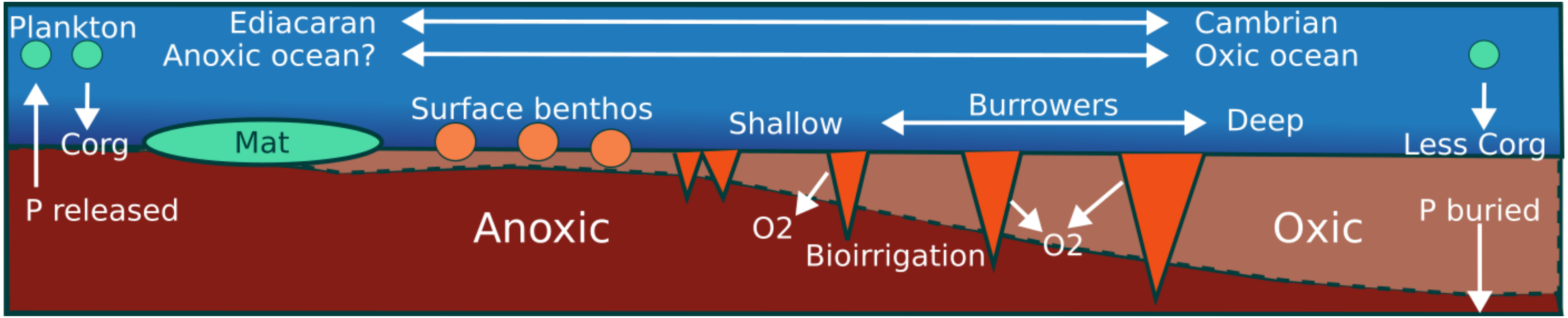



Adapted from Lenton et al 2014; see also Erwin 2011, Logan 1995

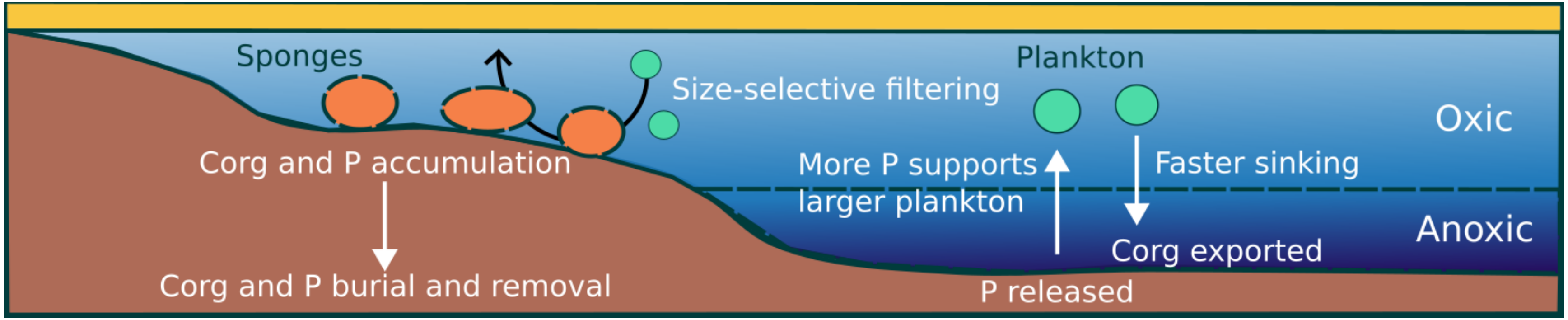
# Group Exercise

How can I use modelling?

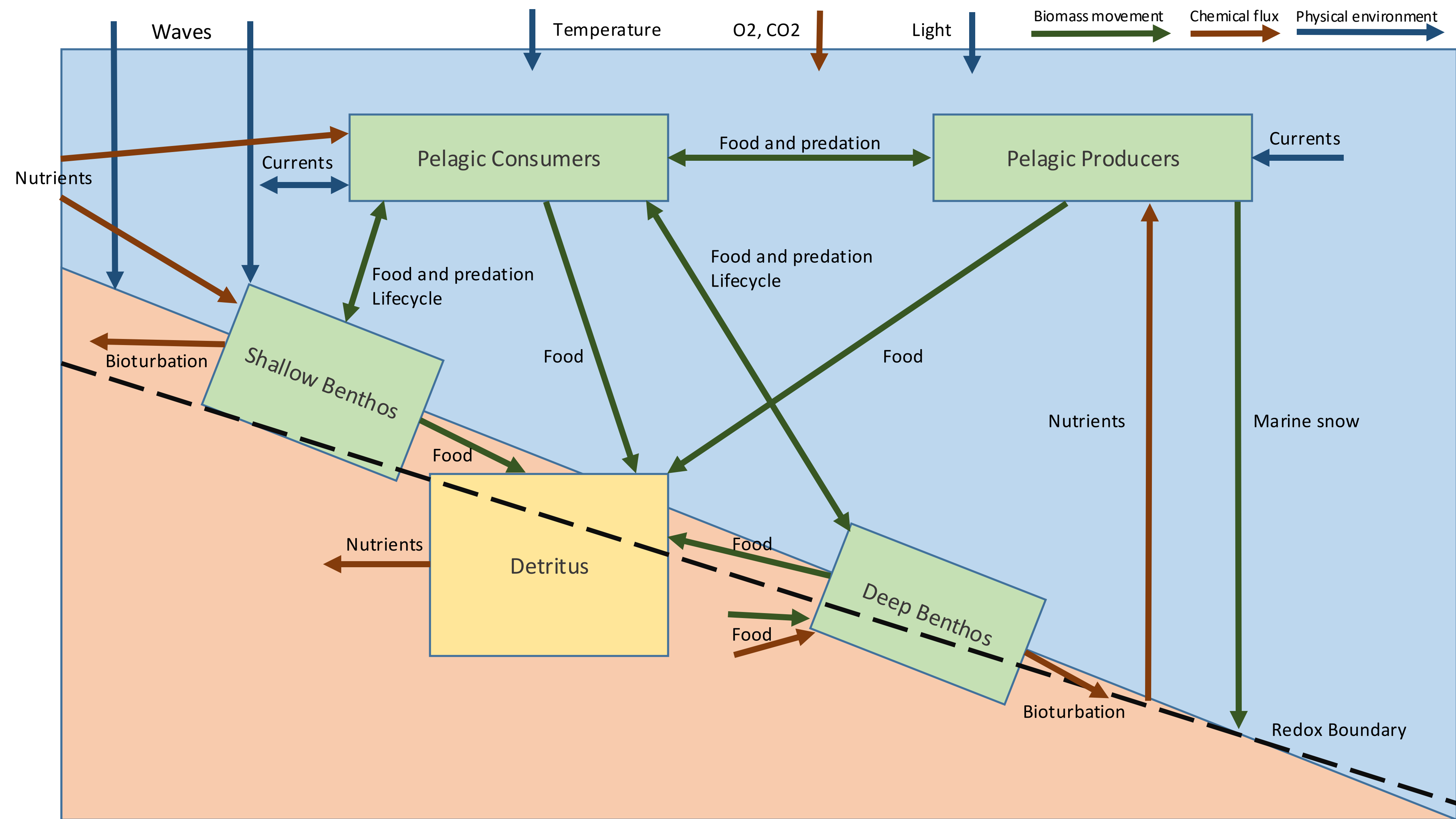
# Cambrian Substrate Revolution




# Sponge size-selective filtering




Adapted from Lenton et al 2014; see also Erwin 2011, Logan 1995


Waves
Temperature
O2, CO2
Light
Biomass movement
Chemical flux
Physical environment
Pelagic Consumers
Food and predation
Pelagic Producers
Currents
Currents
Nutrients
Food and predation
Lifecycle
Food and predation
Lifecycle
Shallow Benthos
Bioturbation
Food
Food
Food
Nutrients
Marine snow
Detritus
Nutrients
Food
Deep Benthos
Food
Bioturbation
Redox Boundary

# Intro to PALEOtoolkit

Stuart Daines

# PALEOtoolkit: scientific needs, approach, and design

- Science drivers for biogeochemical and ecosystem models
    - Processes and scales
    - Piecing together a puzzle
    - Model hierarchies
    - Hypothesis testing
- PALEO approach and design
    - Provide a flexible and extensible way of combining components into systems
    - PALEO components and workflow
- PALEOtoolkit examples
    - Models
    - Comparison to other tools

# Processes and spatial scales

**Individual**

- functional traits
- ecophysiology
- energy and substrate budgets

**Population**

- growth
- death

**Community**

- resource competition
- predation

**Ecosystem**

- energy flow
- material flow
- spatial transport

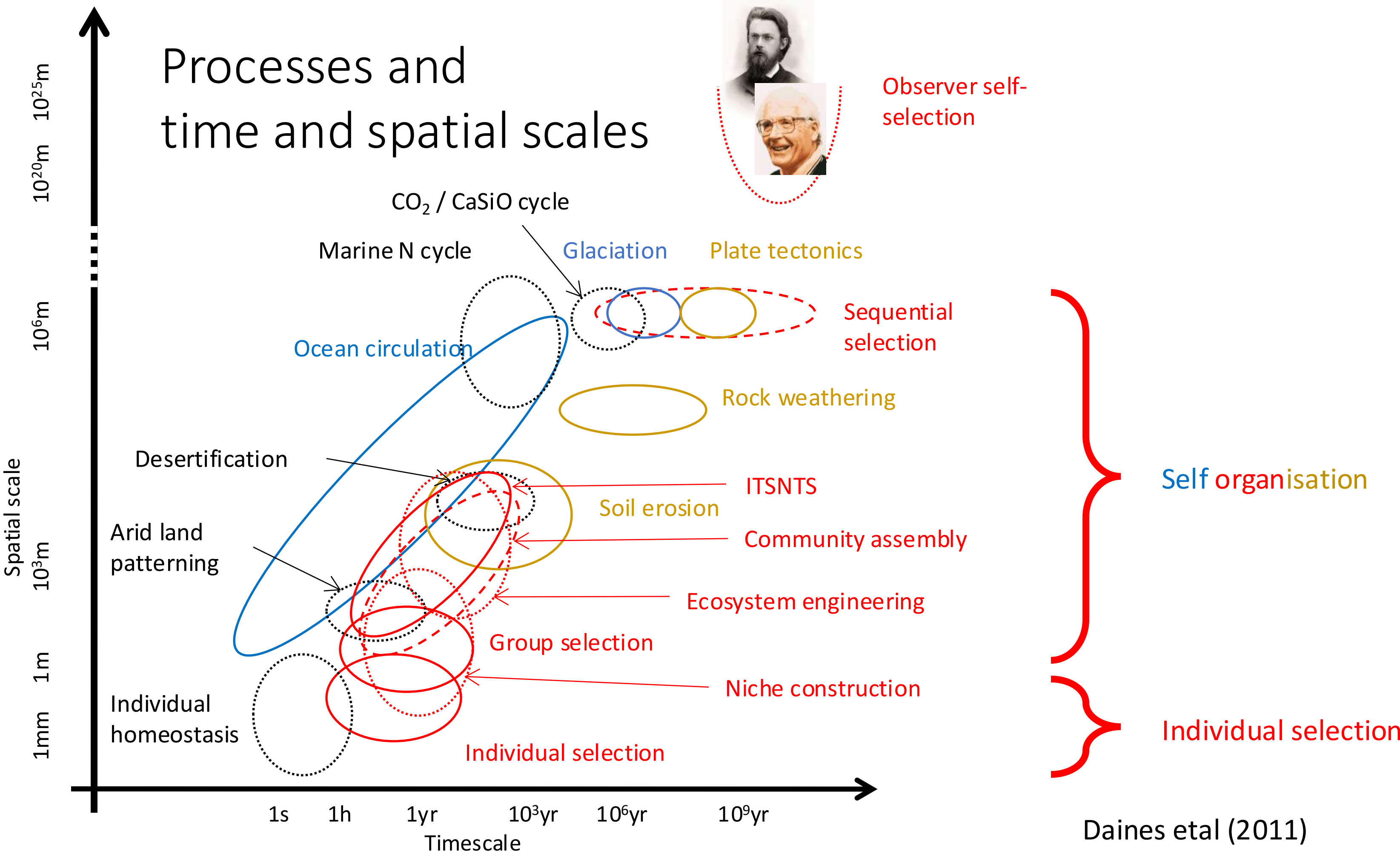
**Biosphere**

- feedbacks
- global BGC

Darwinian natural selection
Evolutionary ecology

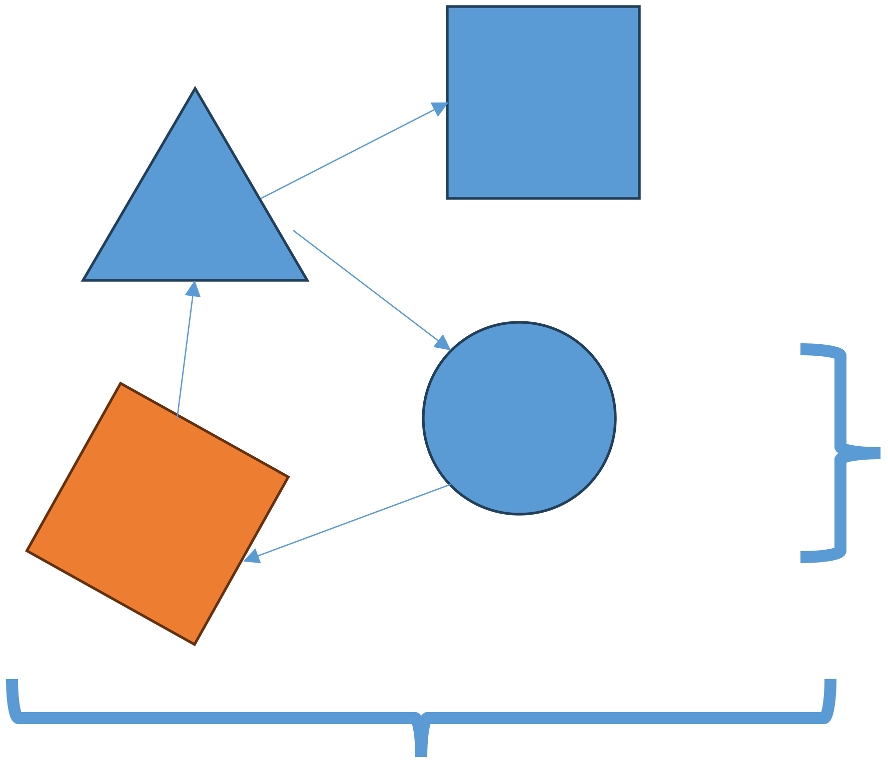
Biosphere dynamics
(Lovelock)

# Processes and time and spatial scales




Daines etal (2011)

# Piecing together a puzzle



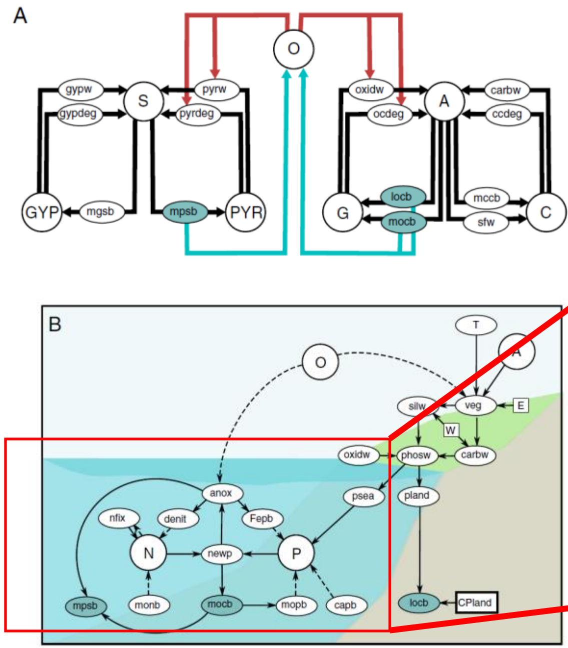
Local (process): constraints observations

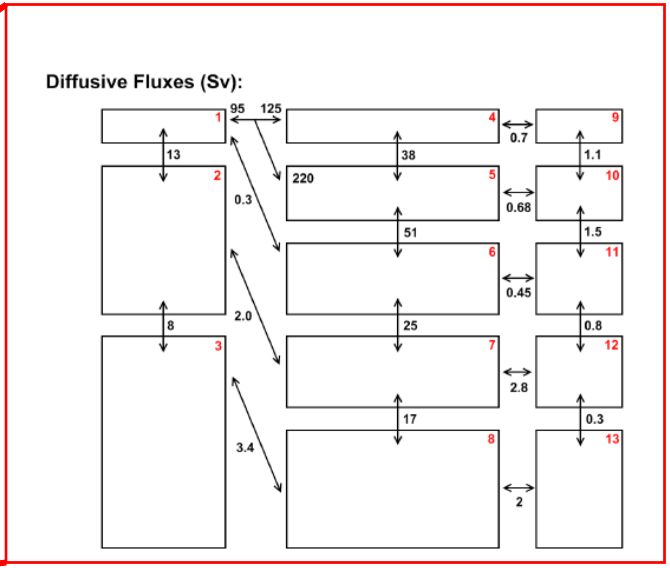
System level: dynamics, observations

# Model hierarchy

COPSE reloaded

PALEO
3 x 1D column ocean (79 box)
Romaniello & Derry (2010)



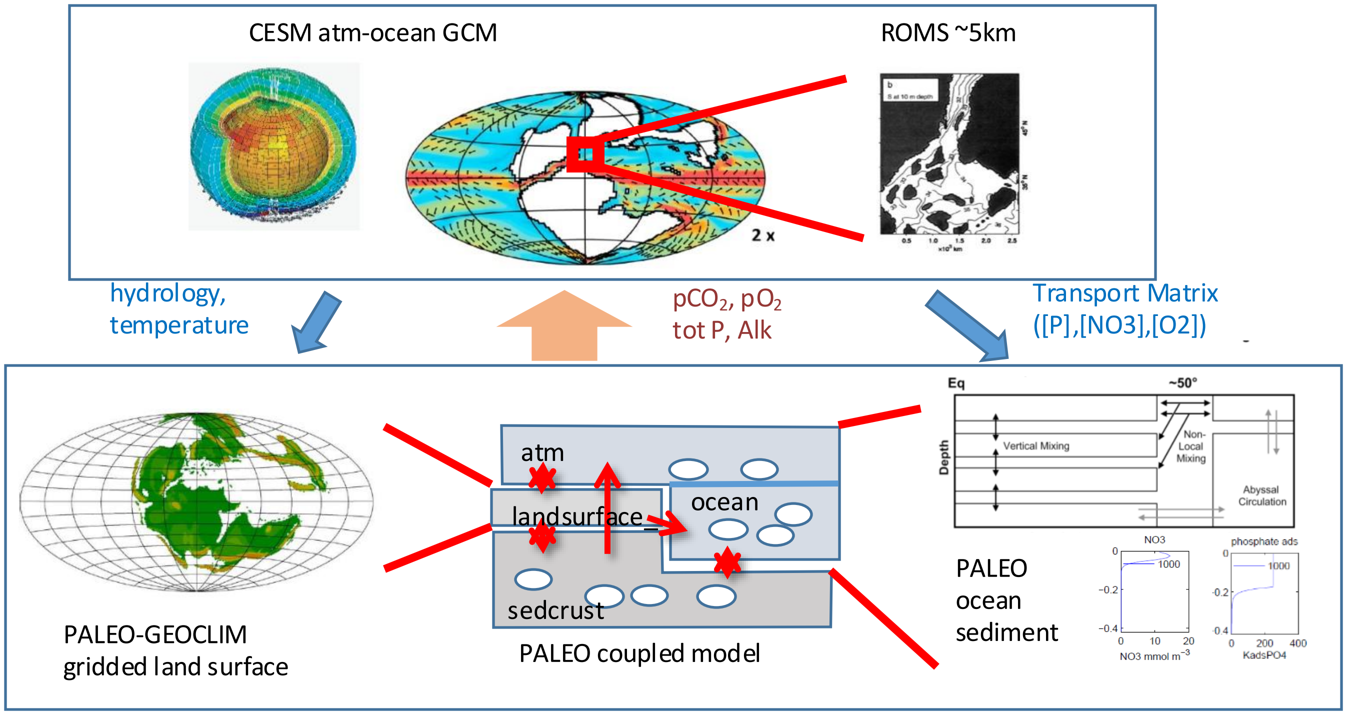




Lenton, Daines, Mills (2018) *Earth Science Reviews*

Clarkson etal (2015) *Science* (using 3-box ocean)

# Model hierarchy exploiting timescale separation:

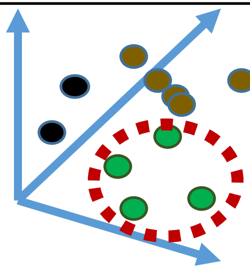
# GCM ‘snapshots’ <-> ‘offline’ long-timescale models



# Hypothesis testing and knowledge generation

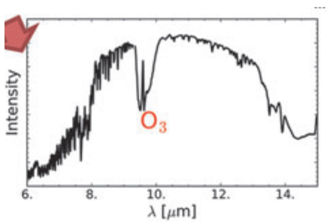
(eg exoplanet biosignatures: cf Lovelock (1965) *Nature*: 'A physical basis for life detection experiments')

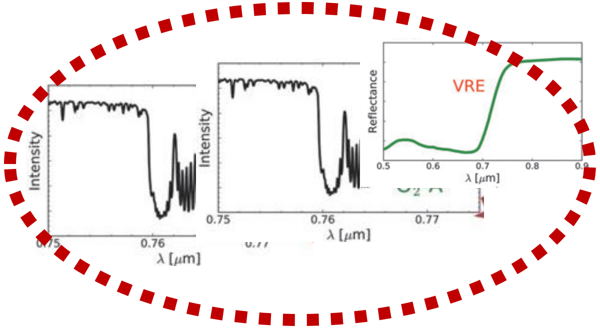
Hypothesis space:
- atm composition
- solid Earth composition
- metabolisms (or no life)
- ...



exo-Earth system models

Predictions
(transmission, reflection, ...)




Observations

Can be formalized:
Catling etal (2018)

# Model needs:

*A flexible way of combining components into systems to make predictions and test hypotheses*

# PALEO approach and design

0D Earth system model:

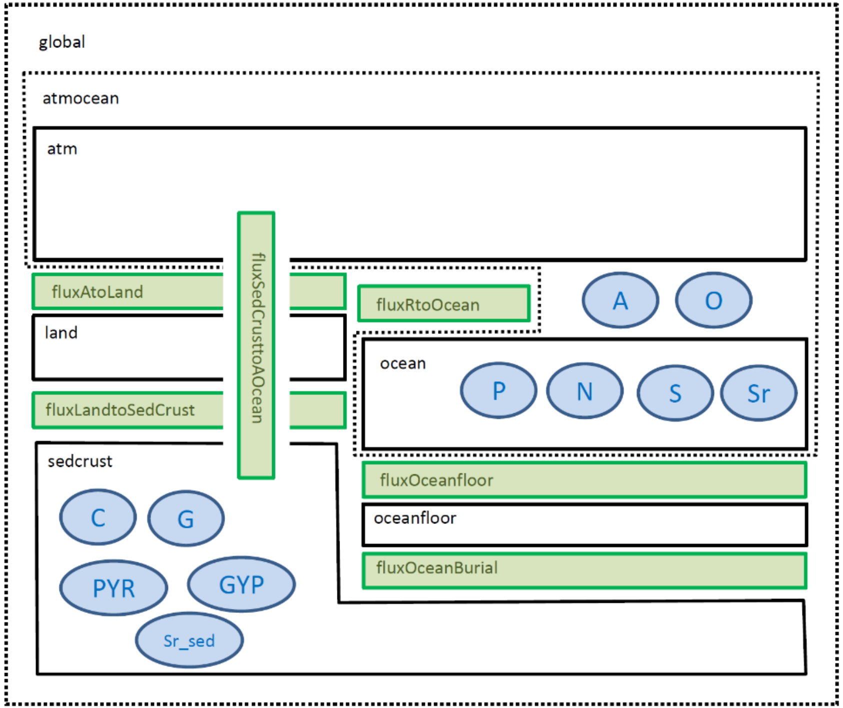
System of ~10 differential equations for Earth surface and sedimentary reservoirs:

eg atmosphere-ocean oxygen:

$$\frac{dO}{dt} = sources\ - sinks$$

ocean phosphorus:

$$\frac{dP}{dt} = sources\ - sinks$$



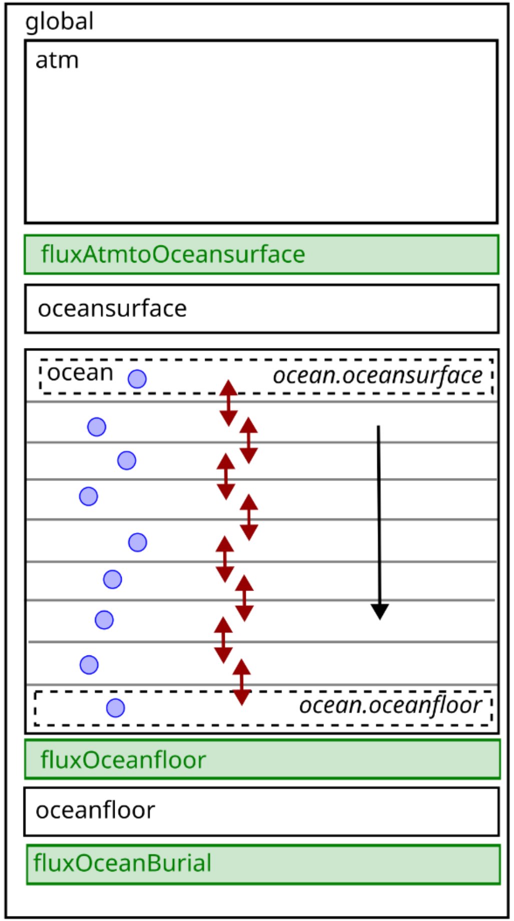
# PALEO approach and design

1D shelf sea water column:
System of ~1000 differential equations for tracers in water column cells (populations, nutrients, ...)

eg phytoplankton population:

$$\frac{dP1\_conc}{dt} = \underbrace{growth}_{\text{per-cell growth rate}} - \underbrace{death}_{\text{per-cell death rate}} + \underbrace{T\ P1\_conc}_{\text{transport by circulation}} + \underbrace{V\ P1\_conc}_{\text{vertical sinking}}$$

Transport is implemented using "transport matrix" representing ocean circulation calculated by an (offline) physical model (eg Khatiwala 2007) GBC



# PALEO represents components as "Reactions"

eg phytoplankton population:

ReactionReservoir: P1_conc (state variable) P1_sms (time derivative)

$$\frac{dP1\_conc}{dt} = \overbrace{\underbrace{growth}_{\substack{\text{Reaction}\\ \text{Phytoplankton}}} - \underbrace{death}_{\substack{\text{Reaction}\\ \text{Zooplankton}}} + \underbrace{T\ P1\_conc}_{\substack{\text{Reaction}\\ \text{OceanTransportColumn}}} + \underbrace{V\ P1\_conc}_{\substack{\text{Reaction}\\ \text{SinkFloat}}}}$$

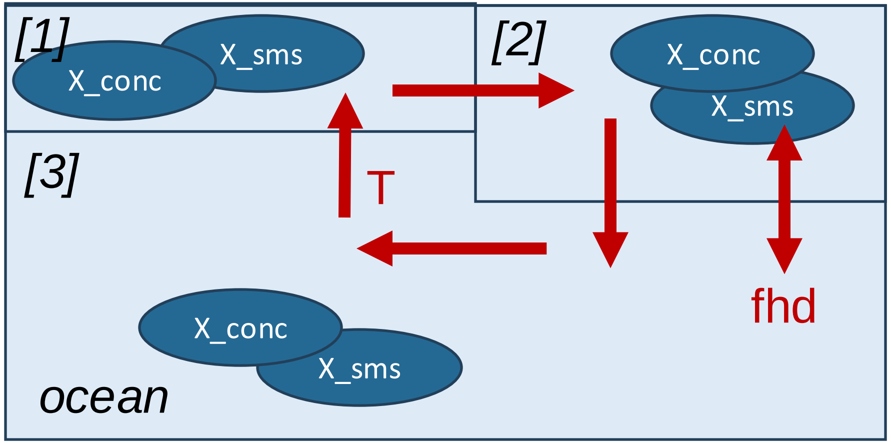
*(most Reactions implement processes that transfer fluxes between reservoirs and appear as terms in ODEs)*

# Minimal example: Reservoir + transport

Sarmiento & Toggweiler (1984) 3-box model of global ocean

ReactionReservoir: X_conc, X_sms

ReactionOceanTransport3box




$$\frac{d}{dt}\begin{pmatrix} X\_conc[1] \\ X\_conc[2] \\ X\_conc[3] \end{pmatrix} = \begin{pmatrix} X\_sms[1] \\ X\_sms[2] \\ X\_sms[3] \end{pmatrix} = \begin{pmatrix} -T & 0 & T \\ T & -(T+fhd) & fhd \\ 0 & T+fhd & -(T+fhd) \end{pmatrix} \begin{pmatrix} X\_conc[1] \\ X\_conc[2] \\ X\_conc[3] \end{pmatrix}$$

# PALEO components and workflow

YAML configuration file

custom Reaction code

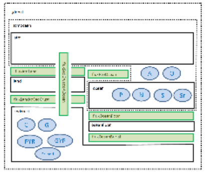
PALEOocean Reaction catalog

PALEOboxes Reaction catalog

PALEOboxes (model coupler)

- create Reactions from catalog
- link variables by name
- set parameters

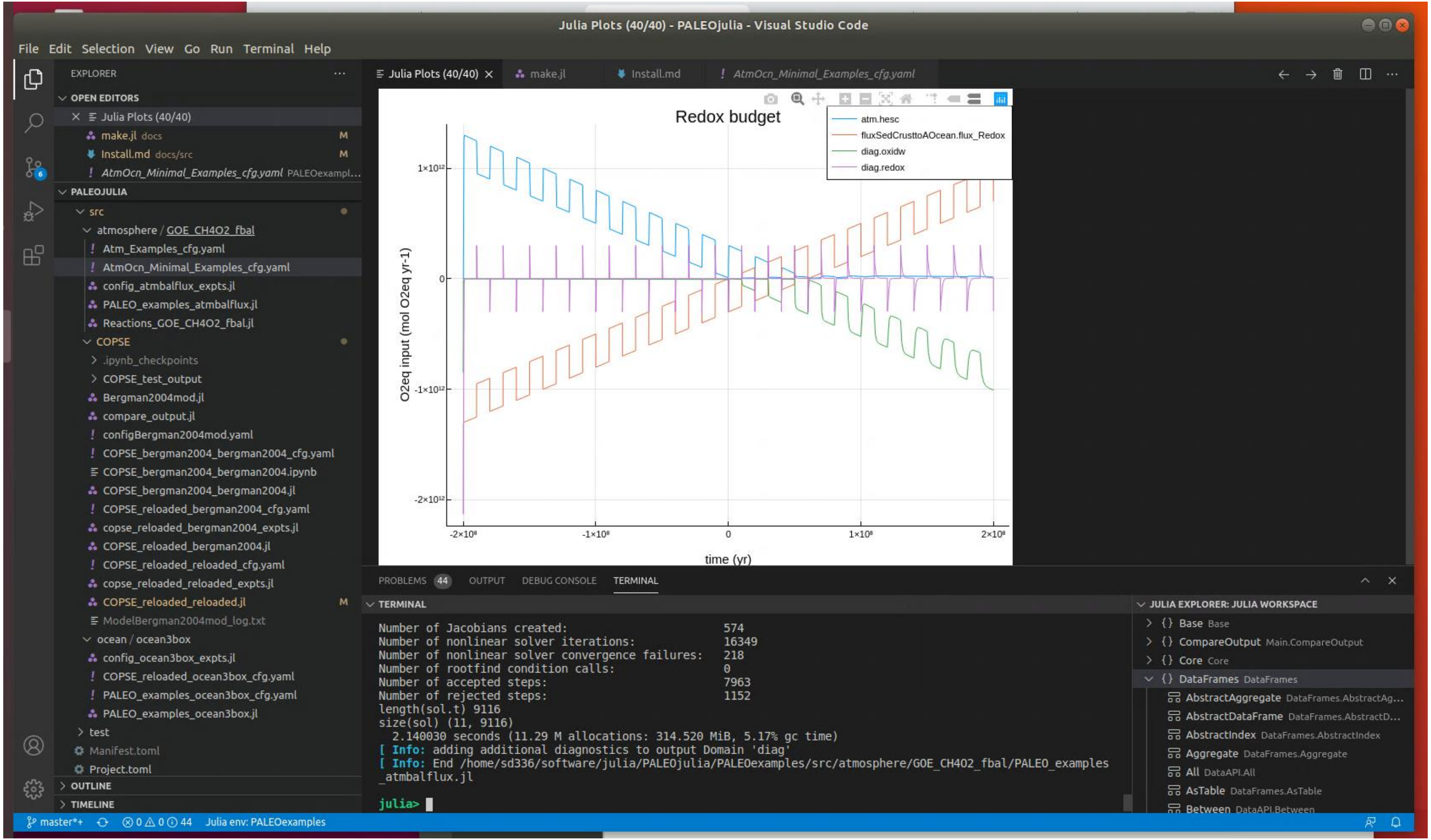
model object



PALEOmodel (solvers, analysis)

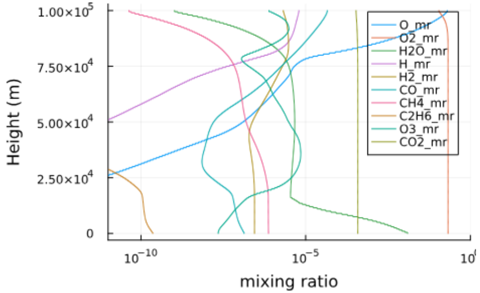
- solve system of ODE
- compare output to datasets
- test parameter sensitivity

output results

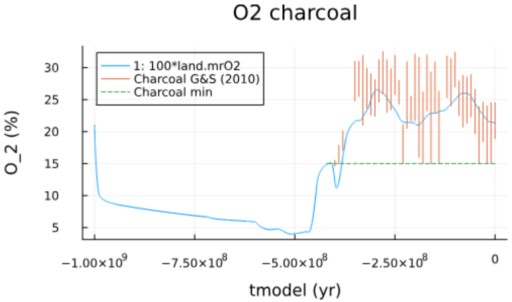

Julia Plots (40/40) - PALEOjulia - Visual Studio Code
File Edit Selection View Go Run Terminal Help
EXPLORER
OPEN EDITORS
Julia Plots (40/40)
make.jl docs M
Install.md docs/src M
AtmOcn_Minimal_Examples_cfg.yaml PALEOexampl...
PALEOJULIA
src
atmosphere / GOE_CH4O2_fbal
Atm_Examples_cfg.yaml
AtmOcn_Minimal_Examples_cfg.yaml
config_atmbalflux_expts.jl
PALEO_examples_atmbalflux.jl
Reactions_GOE_CH4O2_fbal.jl
COPSE
.ipynb_checkpoints
COPSE_test_output
Bergman2004mod.jl
compare_output.jl
configBergman2004mod.yaml
COPSE_bergman2004_bergman2004_cfg.yaml
COPSE_bergman2004_bergman2004.ipynb
COPSE_bergman2004_bergman2004.jl
COPSE_reloaded_bergman2004_cfg.yaml
copse_reloaded_bergman2004_expts.jl
COPSE_reloaded_bergman2004.jl
COPSE_reloaded_reloaded_cfg.yaml
copse_reloaded_reloaded_expts.jl
COPSE_reloaded_reloaded.jl M
ModelBergman2004mod_log.txt
ocean / ocean3box
config_ocean3box_expts.jl
COPSE_reloaded_ocean3box_cfg.yaml
PALEO_examples_ocean3box_cfg.yaml
PALEO_examples_ocean3box.jl
test
Manifest.toml
Project.toml
OUTLINE
TIMELINE
Julia Plots (40/40)
make.jl
Install.md
AtmOcn_Minimal_Examples_cfg.yaml
Redox budget
atm.hesc
fluxSedCrusttoAOcean.flux_Redox
diag.oxidw
diag.redox
O2eq input (mol O2eq yr-1)
1×10¹²
0
-1×10¹²
-2×10¹²
-2×10⁸
-1×10⁸
0
1×10⁸
2×10⁸
time (yr)
PROBLEMS 44
OUTPUT
DEBUG CONSOLE
TERMINAL
TERMINAL
Number of Jacobians created: 574
Number of nonlinear solver iterations: 16349
Number of nonlinear solver convergence failures: 218
Number of rootfind condition calls: 0
Number of accepted steps: 7963
Number of rejected steps: 1152
length(sol.t) 9116
size(sol) (11, 9116)
2.140030 seconds (11.29 M allocations: 314.520 MiB, 5.17% gc time)
[ Info: adding additional diagnostics to output Domain 'diag'
[ Info: End /home/sd336/software/julia/PALEOjulia/PALEOexamples/src/atmosphere/GOE_CH4O2_fbal/PALEO_examples_atmbalflux.jl
julia>
JULIA EXPLORER: JULIA WORKSPACE
{} Base Base
{} CompareOutput Main.CompareOutput
{} Core Core
{} DataFrames DataFrames
AbstractAggregate DataFrames.AbstractAg...
AbstractDataFrame DataFrames.AbstractD...
AbstractIndex DataFrames.AbstractIndex
Aggregate DataFrames.Aggregate
All DataAPI.All
AsTable DataFrames.AsTable
Between DataAPI.Between
master*+
0 0 44
Julia env: PALEOexamples

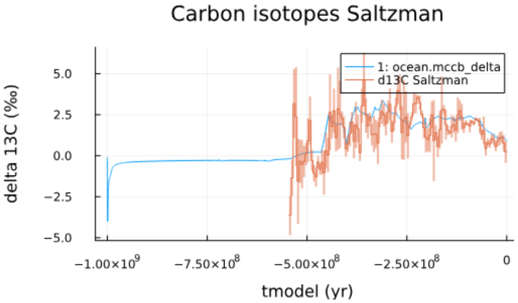
# PALEOtoolkit models

PALEOatmosphere.jl
(photo)chemistry

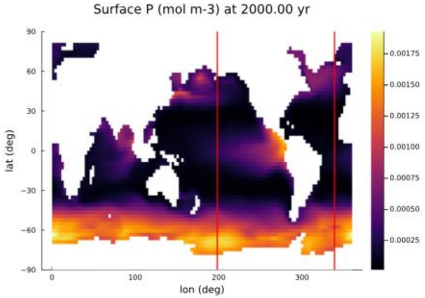


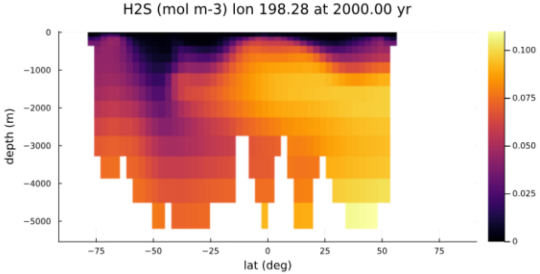

PALEOcopse.jl
BGC Earth history




PALEOocean.jl
MITgcm2.8 deg



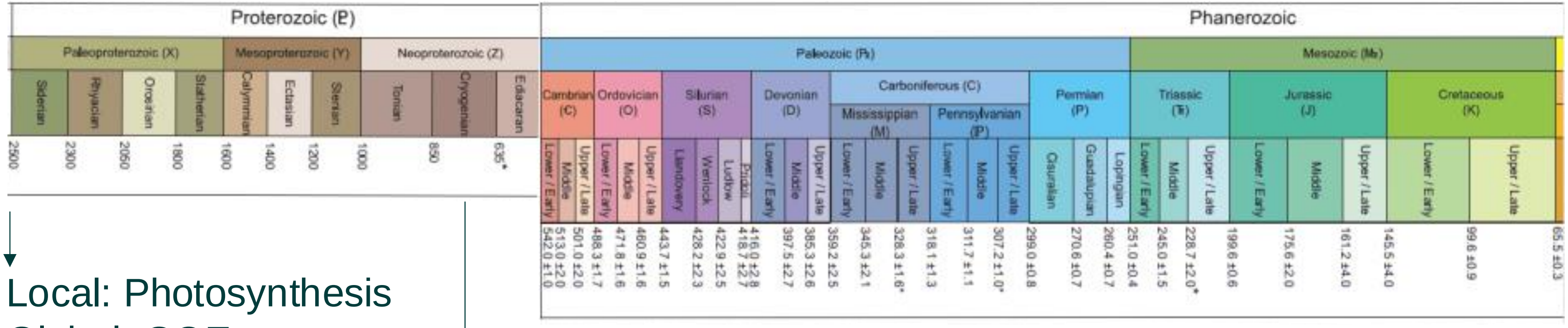

PALEOsediment.jl (reaction-transport)

# Comparison of model tools

- PALEO: flexible, extensible, hierarchical
  - Julia: fast ! (100x python or R)
  - automatic differentiation (Jacobians, MCMC ...)
  - Julia package ecosystem ! SciML
- GEOCLIM, SCION: biogeochemistry + offline climate
- GENIE: monolithic ESM
- FABM (marine ecosystem): GCM focused
- R ReactTran (sediments), ...
- ATMOS (atmospheric chemistry), ...

# Tutorial 1

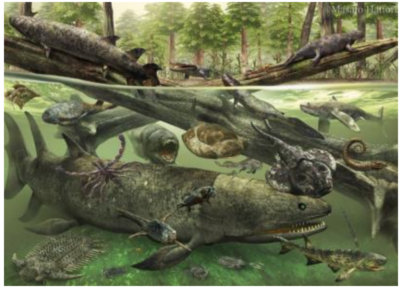
Group walk-through of NP model




Local: Photosynthesis
Global: GOE

Local: Sponge evolution
Global: Redox change

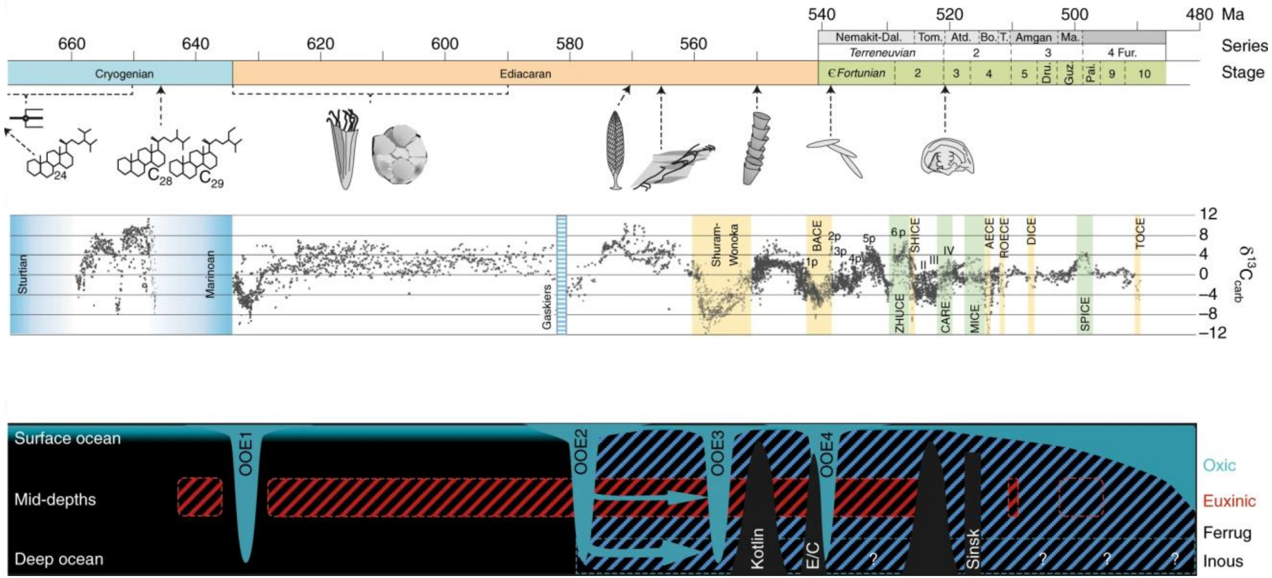
Local: Plant evolution
Global: LDME



Local: Biotic innovations
Global: Redox change

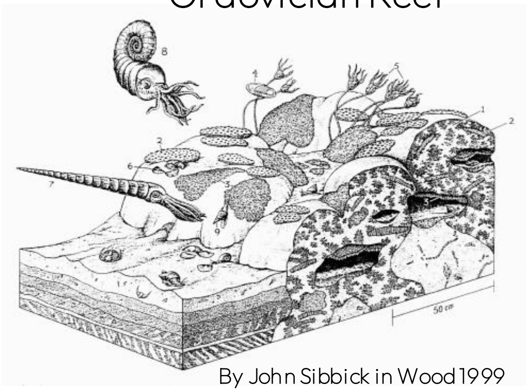
# Ancient Shelf Seas

- Shelf sea setting and redox state drive evolutionary events
- Correlating sedimentary, geochemical and fossil data insufficient



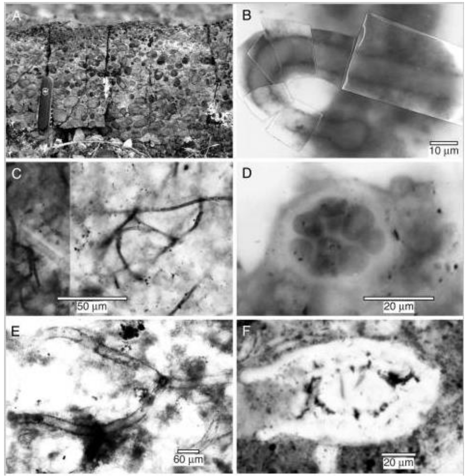

Wood et al 2019

## Ordovician Reef



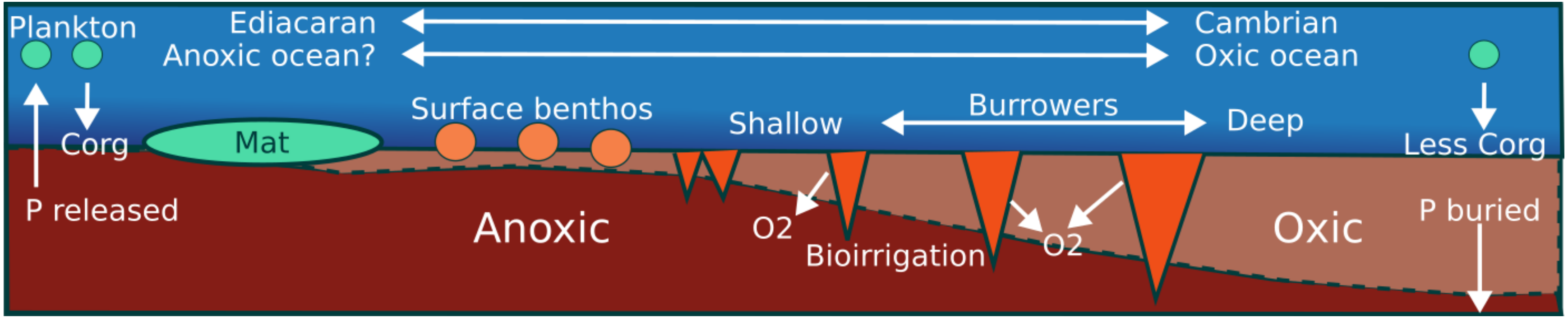
By John Sibbick in Wood 1999

## Neoproterozoic Microfossils

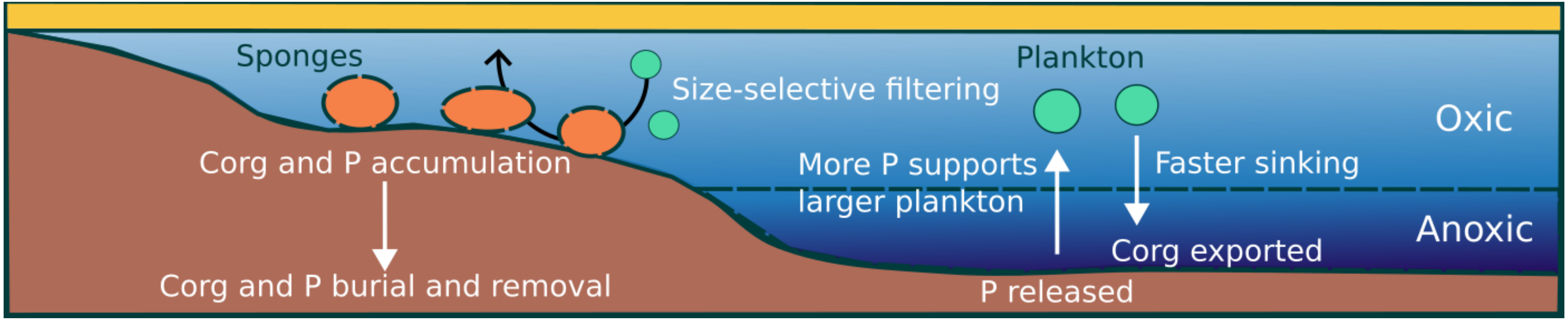



Corsetti et al 2003

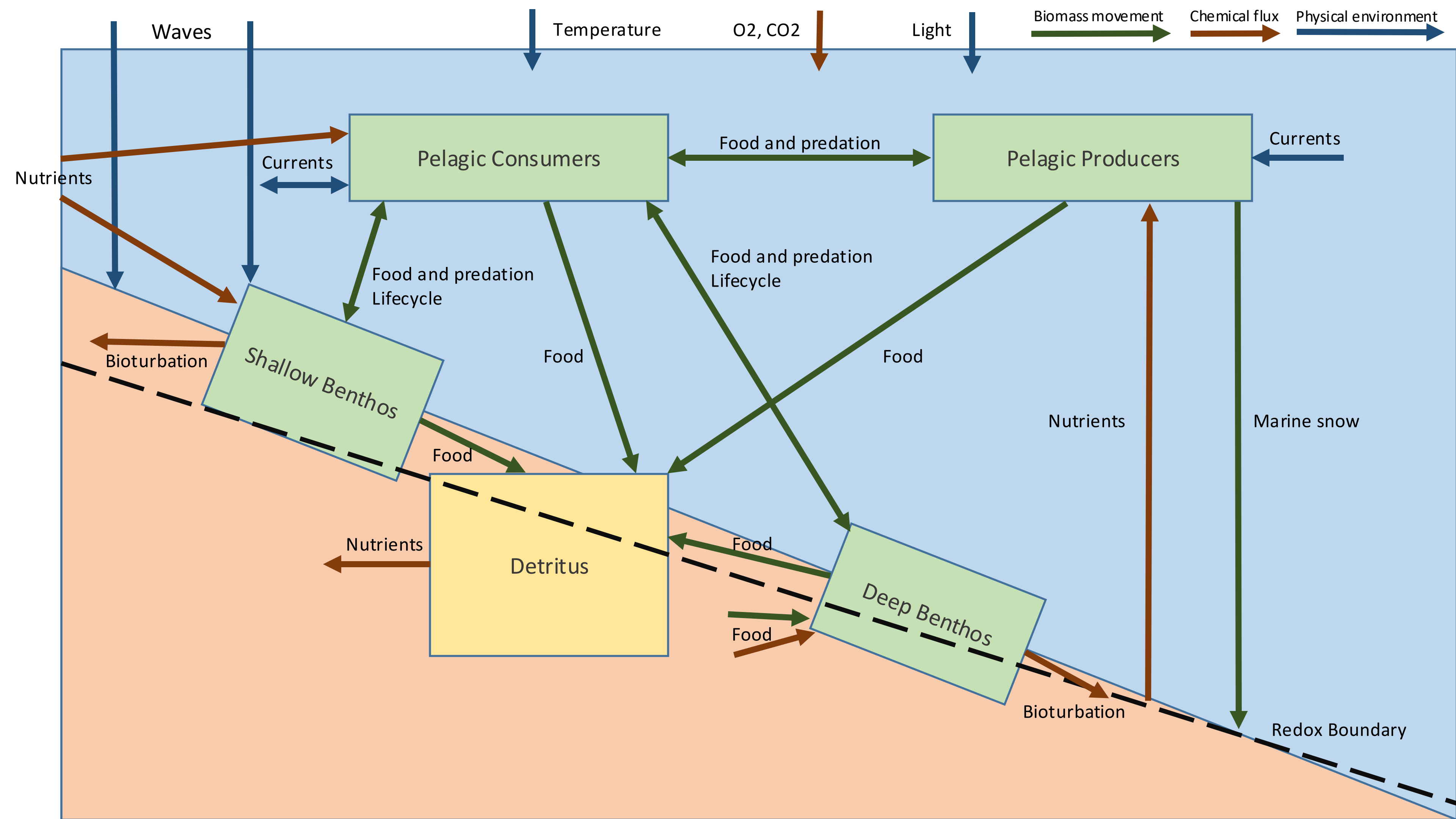
# Cambrian Substrate Revolution



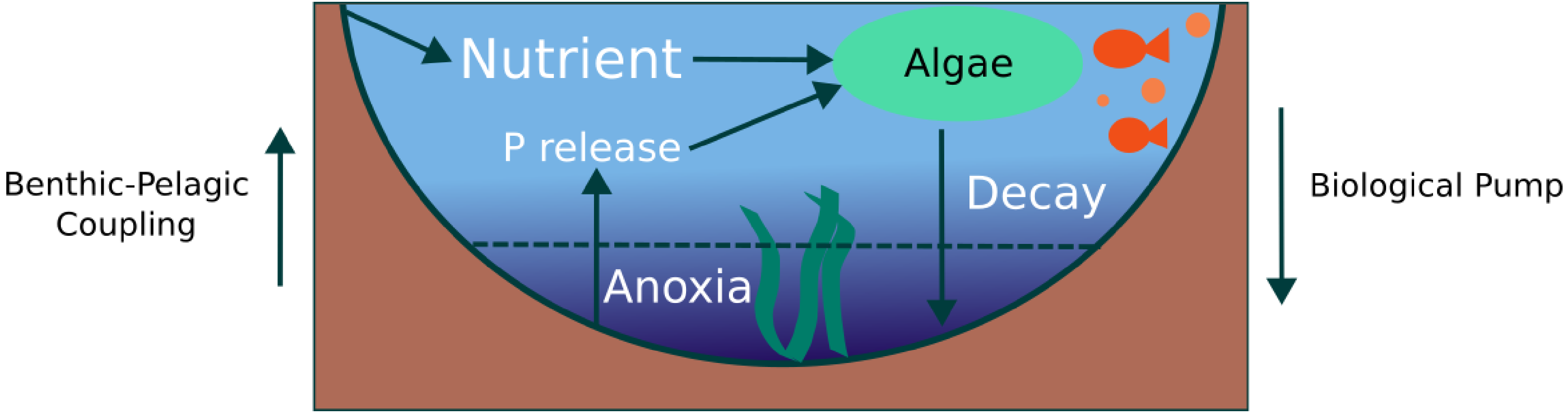

# Sponge size-selective filtering




Adapted from Lenton et al 2014; see also Erwin 2011, Logan 1995

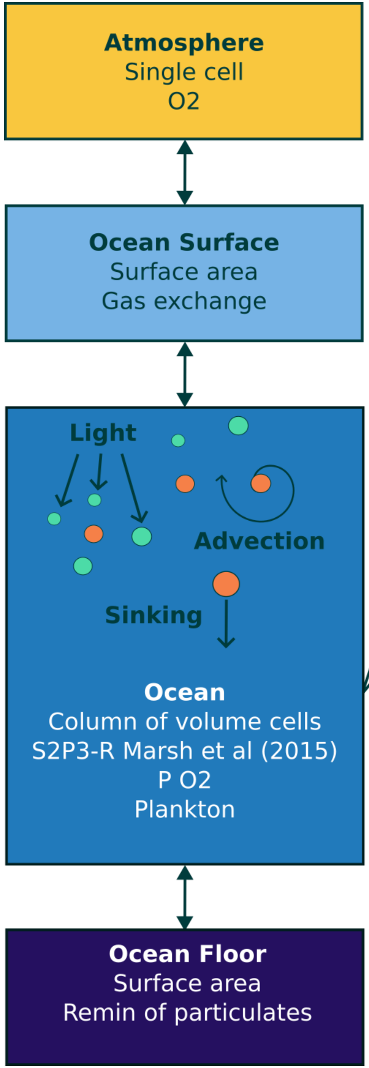

Waves
Temperature
O2, CO2
Light
Biomass movement
Chemical flux
Physical environment
Nutrients
Currents
Pelagic Consumers
Food and predation
Pelagic Producers
Currents
Food and predation
Lifecycle
Food
Food and predation
Lifecycle
Food
Bioturbation
Shallow Benthos
Food
Nutrients
Marine snow
Nutrients
Detritus
Food
Deep Benthos
Food
Bioturbation
Redox Boundary

# Dynamic Feedbacks are Critical

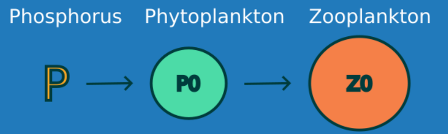



Bidirectional effects in eutrophication

# 1D Shelf Sea Column Model




**Plankton**




Assimilate P as in Armstrong (1994):

For nutrient assimilation in autotrophs:

$$\frac{\partial P_i}{\partial dt} = P_i(\mu_{max,i}\frac{S}{K_{S,i}+S} - \lambda_i - Z_i\frac{H_{max,i}}{K_{P,i}})$$

For feeding in heterotrophs:

$$\frac{\partial Z_i}{\partial dt} = Z_i(\gamma_i\frac{H_{max,i}}{K_{P,i}}P_i - \delta_i)$$

$S$ nutrient concentration, mmol N $m^{-3}$; $P_i$ population of plankton, mmol N $m^{-3}$; $Z_i$ population of zooplankton, mmol N $m^{-3}$; $\mu_{max,i}$ maximum phytoplankton, growth rate $d^{-1}$; $\lambda_i$ plankton growth, mmol N $m^{-3}$; $H_{max,i}$ maximum zooplankton harvest rate, mmol P $d^{-1}$; $K_{P,i}$ zooplankton full-saturation constant for harvest and growth, mmol P $m^{-3}$; $\gamma_i$ zooplankton growth efficiency, dimensionless; $\delta_i$ intrinsic zooplankton loss rate, $d^{-1}$

- **Spatial resolution** – 1D enough for redoxcline and mass budgets
- **Dynamics** – S2P3-R column model enough for some physical transport, light/temp profile
- **Elements** – P and O enough to capture basic redox and nutrient relationship
- **"Biological pump"** – plankton food web base

# NPZ Model Output

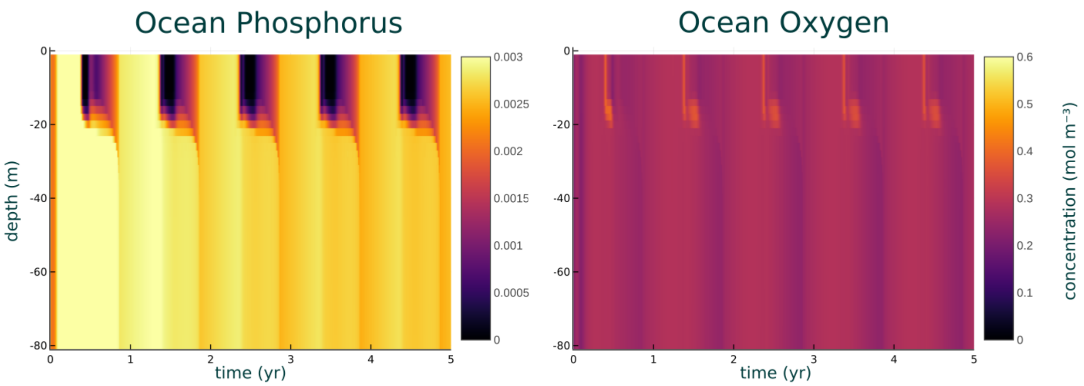



**Figure 1a,b:** As the seasonal solar energy flux increases, the light-dependent phytoplankton begin to consume phosphorous and produce oxygen.

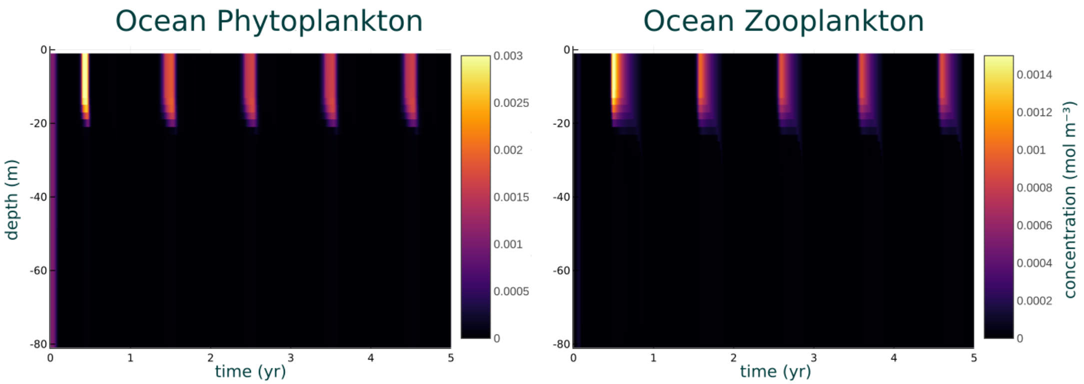



**Figure 2a,b:** Zooplankton feed on the phytoplankton, reducing their population and growing their own.

$$\frac{\partial Z_i}{\partial dt} = Z_i(\gamma_i \frac{H_{max,i}}{K_{P,i}} P_i - \delta_i)$$

- Gamma: 0.4, units=dimensionless; assimilation efficiency
- $H_{max}$: 1.4, units=$d^{-1}$; max harvest rate
- $K_p$: 2e-3, units=mol $m^{-3}$; full-saturation constant

# Closing

Final reflections

# Reflection on Workshop Goals

Participants will leave with...

- an understanding of how to use models to generate knowledge;
- a firmer sense of what modelling is, if new to modelling;
- the ability to define models to test hypotheses about the Earth system;
- awareness of how PALEOtoolkit can be used to test Earth system models;
- confidence engaging with the mathematics and code used for models.