Leveraging microbial structure data to inform traits-based modeling, an approach based on microbial community segmentation.

Emelia J. Chamberlain[1], Hyewon Kim[2], Scott Doney[2,3], Jeff Bowman[1,4]

[1]Integrative Oceanography Division, Scripps Institution of Oceanography - La Jolla, CA USA
[2]Department of Environmental Sciences, University of Virginia - Charlottesville, VA USA
[3]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA USA
[4]Center for Microbiome Innovation, University of California San Diego - La Jolla, CA USA

I. Introduction

Numerical modeling is a critical method for understanding ecosystem processes. However, current approaches are typically not informed by microbial diversity data due to its high dimensionality. This creates a discrepancy between the observed and modeled biological complexity that over-simplifies the dynamics and function of the microbial communities in marine ecosystems. Here we describe a strategy for incorporation of microbial community structure data in the Regional Test-bed Model, a 1-D data assimilative ecosystem model with parameters optimized for the pelagic western Antarctic Peninsula (WAP) near Palmer Station.
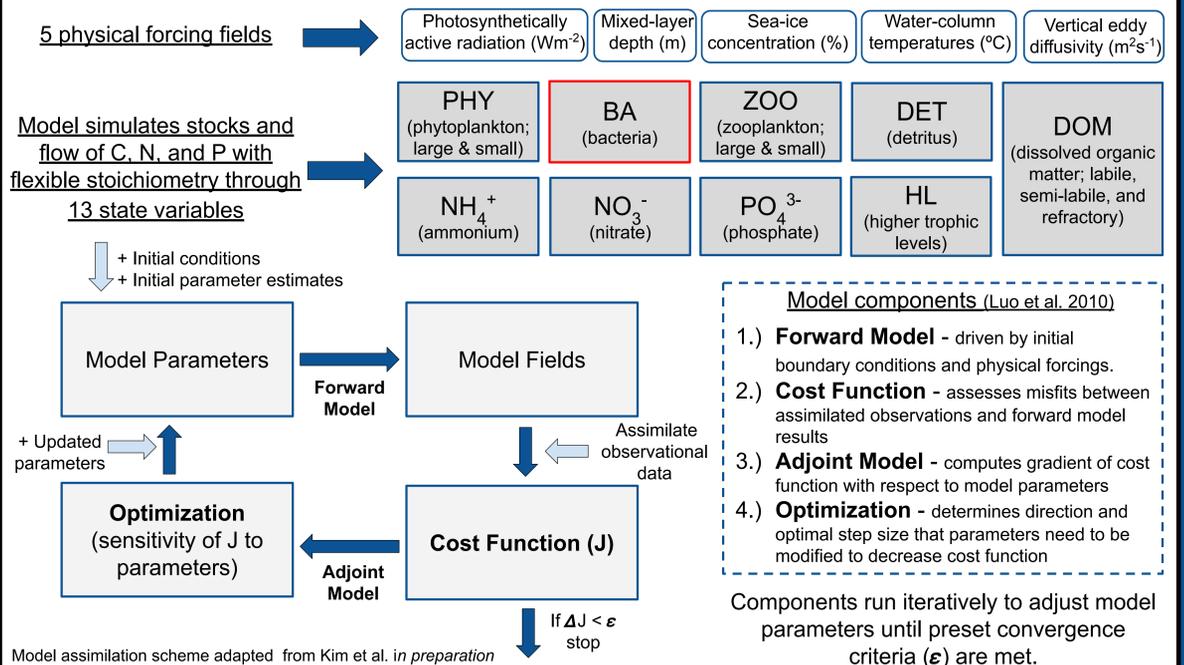
II. Segmentation of Microbial Community Structure

Bowman et al. (2017) introduced a technique to "segment" the microbial community into functionally coherent units ("modes") that can be described by a single categorical variable. This categorical variable reflects the key genetic traits of the microbial community. An analysis of 16S rRNA gene data from a 5-year time-series from the WAP was described by 8 recurrent modes in the bacterial community. Each point in the time series is associated with a specific mode based on microbial community relative abundance. Estimates of genome parameters estimated for each mode showed logical temporal trends in 16S rRNA gene copy number, genome size, and GC content. From these genetic traits we can make reasonable estimates of physiology (e.g., respiration, bacterial growth efficiency, cell size).
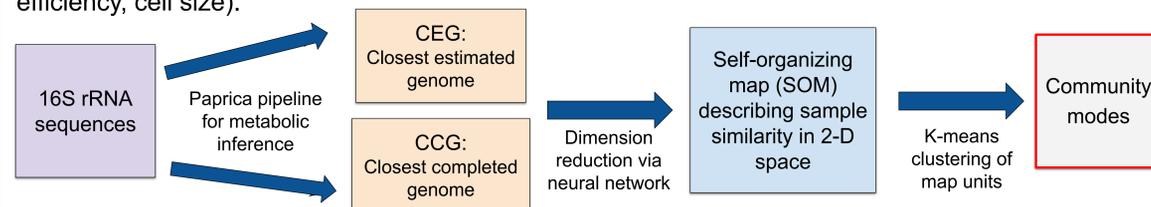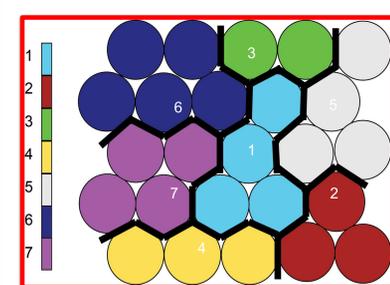

III. Palmer Regional Test-bed Model (Kim et al. *in preparation*)

1-Dimensional data assimilative marine ecosystem model which utilizes a *variational adjoint scheme* (Lawson et al. 1995) to optimize biological model parameters. Built from existing regional test-bed models from other long term time series (Luo et al. 2010, Friedrichs et al. 2006).

5 physical forcing fields → Photosynthetically active radiation (Wm^{-2}) | Mixed-layer depth (m) | Sea-ice concentration (%) | Water-column temperatures (^{o}C) | Vertical eddy diffusivity (m^{2}s^{-1})

Model simulates stocks and flow of C, N, and P with flexible stoichiometry through 13 state variables →

PHY (phytoplankton; large & small) | BA (bacteria) | ZOO (zooplankton; large & small) | DET (detritus) | DOM (dissolved organic matter; labile, semi-labile, and refractory)

NH_4^+ (ammonium) | NO_3^- (nitrate) | PO_4^{3-} (phosphate) | HL (higher trophic levels)

+ Initial conditions
+ Initial parameter estimates

Model Parameters → **Forward Model** → Model Fields

+ Updated parameters

Optimization (sensitivity of J to parameters) → **Adjoint Model** → Cost Function (J)

Assimilate observational data

If $\Delta J < \varepsilon$ stop

Model assimilation scheme adapted from Kim et al. *in preparation*

Model components (Luo et al. 2010)

1.) **Forward Model** - driven by initial boundary conditions and physical forcings.
2.) **Cost Function** - assesses misfits between assimilated observations and forward model results.
3.) **Adjoint Model** - computes gradient of cost function with respect to model parameters
4.) **Optimization** - determines direction and optimal step size that parameters need to be modified to decrease cost function

Components run iteratively to adjust model parameters until preset convergence criteria (ε) are met.

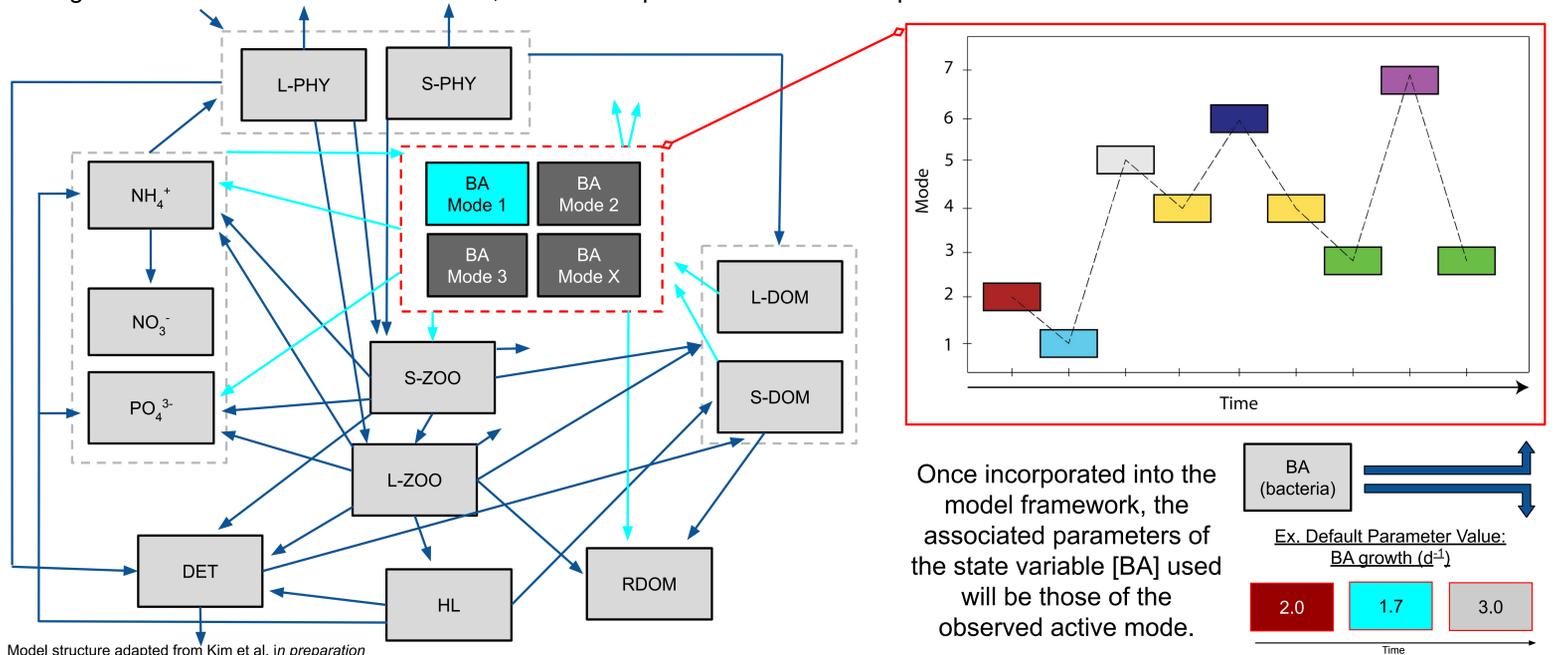## IV. Unifying trait-based microbial community segmentation and numerical modeling

Each community mode has its own functional genetic traits from which we can make reasonable estimates of model parameters

- Maximum BA growth rate
- BA active respiration rate vs. Production
- Bacteria extra semi labile DOC excretion rate
- BA inorganic nutrients regeneration rate
- BA refractory DOC production rate
- BA selection strength on SDOM
- BA basal respiration rate - min/max
- BA mortality rate
- Half-saturation density of BA for grazing

Using a model hind-cast of the time series, we can compare data assimilation parameter estimates with estimates from observed mode.

Ex: Physiology	Mode 2	Mode 6
BA growth efficiency	0.06	0.32
BA active respiration	5 mmol/m^{3}d	10 mmol/m^{3}d
BA carbon biomass	0.74 mmol/m^3	1.20 mmol/m^3

Provides realistic initial conditions and a useful "check" against parameter optimization with data assimilation.

Model structure adapted from Kim et al. *in preparation*

Once incorporated into the model framework, the associated parameters of the state variable [BA] used will be those of the observed active mode.

Ex. Default Parameter Value: BA growth (d^{-1})

2.0 | 1.7 | 3.0

V. Conclusion

By reducing the dimensionality of microbial diversity data to enable a correlation between community structure and function we are able to better incorporate microbial dynamics into predictive models. Ultimately, we expect this approach to improve the fidelity of traits-based ecosystem models, leading to a better understanding of complex ecosystem processes.

References:
- Bowman J.S., L.A. Amaral-Zettler, J.J. Rich, C.M. Luria, H.W. Ducklow. (2017). Bacterial community segmentation facilitates the prediction of ecosystem function along the coast of the western Antarctic Peninsula. The ISME Journal. 11: 1460-1471
- Friedrichs MAM, Hood RR, Wiggert JD (2006). Ecosystem model complexity versus physical forcing: quantification of their relative impact with assimilated Arabian Sea data. Deep-Sea Res. II, 53:576-600.
- Kim et al. *in preparation*
- Lawson LM, Spitz YH, Hofmann EE, Long RB (1995). A data assimilation technique applied to a predator-prey model. Bull. Math. Biol., 57:593-617.
- Luo, Y.W., M.A.M. Friedrichs, S.C. Doney, M.J. Church, H.W. Ducklow. (2010). Oceanic heterotrophic bacterial nutrition by semilabile DOM as revealed by data assimilative modeling. Aquatic Microbial Ecology. 60: 273-287.