Supplementary Information:

A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes

Joshua S. Weitz[1], School of Biology and Physics, Georgia Institute of Technology, Atlanta, GA, USA, **Charles A. Stock**, Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, NJ, USA, **Steven W. Wilhelm**, Department of Microbiology, University of Tennessee, Knoxville, TN, USA, **Lydia Bourouiba**, Department of Applied Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA, **Alison Buchan**, Department of Microbiology, University of Tennessee, Knoxville, TN, USA, **Maureen L. Coleman**, Department of Geosciences, University of Chicago, Chicago, IL, USA, **Michael J. Follows**, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, **Jed A. Fuhrman**, Department of Biological Sciences,University of Southern California, Los Angeles, CA, USA, **Luis F. Jover**, School of Physics, Georgia Institute of Technology, Atlanta, GA, USA, **Jay T. Lennon**, Department of Biology, Indiana University, Bloomington, IN, USA, **Mathias Middelboe**, Marine Biological Section, University of Copenhagen, Copenhagen, Denmark, **Derek L. Sonderegger**, Department of Mathematics, Northern Arizona University, Flagstaff, AZ, USA, **Curtis A. Suttle**, Department of Earth and Ocean Sciences, Department of Botany and Department of Immunology and Microbiology, University of British Columbia, Vancouver, British Columbia, Canada, **Bradford P. Taylor**, School of Physics, Georgia Institute of Technology, Atlanta, GA, USA, **T. Frede Thingstad**, Department of Biology, University of Bergen, Bergen, Norway, **William H. Wilson**, Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA; current address: Plymouth Marine Laboratory, Plymouth, UK and **K. Eric Wommack**, Delaware Biotechnology Institute, University of Delaware, Newark, DE, USA

1 NPHZ-V model of microbial ecosystem dynamics

1.1 Model with viruses

We propose the following systems of equations to represent the dynamic changes of biotic populations, including heterotrophs (H), cynanobacteria (C), eukaryotic autotrophs (E), zooplankton (Z) and viruses (V_i where $i = H, C$, and E), along with organic and inorganic nutrients (x_{on} and x_{in} respectively). The system of equations (S1-S9) are nonlinear, coupled ODEs. Eq. (S10) makes explicit the export from the system to higher trophic levels. We use units of particles/L for all populations and μmol/L for nutrient concentrations. Hence, conversion factors, q, denote the equivalent nitrogen content of cells. Definition of parameters are in Table S1. Parameters include those associated with interactions and with the nutrient content of each biotic population on a per-cell basis.

[1]Email: jsweitz@gatech.edu

$$\text{Heterotrophs: } \dot{H} = \overbrace{\frac{\mu_H H x_{on}}{x_{on}+K_{on}}}^{\text{H growth}} - \overbrace{\phi_{VH} H V_H}^{\text{viral lysis}} - \overbrace{\psi_{ZH} H Z}^{\text{grazing}} - \overbrace{m_{on,H} H}^{\text{organic loss}} - \overbrace{m_{in,H} H}^{\text{respiration}} \tag{S1}$$

$$\text{Cyanobacteria: } \dot{C} = \overbrace{\frac{\mu_C C x_{in}}{x_{in}+K_{in,C}}}^{\text{C growth}} - \overbrace{\phi_{VC} C V_C}^{\text{viral lysis}} - \overbrace{\psi_{ZC} C Z}^{\text{grazing}} - \overbrace{m_{on,C} C}^{\text{organic loss}} - \overbrace{m_{in,C} C}^{\text{respiration}} \tag{S2}$$

$$\text{Euk. autos: } \dot{E} = \overbrace{\frac{\mu_E E x_{in}}{x_{in}+K_{in,E}}}^{\text{E growth}} - \overbrace{\phi_{VE} E V_E}^{\text{viral lysis}} - \overbrace{\psi_{ZE} E Z}^{\text{grazing}} - \overbrace{m_{on,E} E}^{\text{organic loss}} - \overbrace{m_{in,E} E}^{\text{respiration}} \tag{S3}$$

$$\text{Zooplankton } \dot{Z} = \overbrace{p_g \left(\frac{q_H}{q_Z} \psi_{ZH} H Z + \frac{q_C}{q_Z} \psi_{ZC} C Z + \frac{q_E}{q_Z} \psi_{ZE} E Z \right)}^{\text{grazing}} - \overbrace{m_Z Z}^{\text{respiration}} - \overbrace{m_{ZP} Z^2}^{\text{consumption}} \tag{S4}$$

$$\text{Viruses of H: } \dot{V_H} = \overbrace{\beta_H \phi_{VH} H V_H}^{\text{lysis}} - \overbrace{m_{VH} V_H}^{\text{decay}} \tag{S5}$$

$$\text{Viruses of C: } \dot{V_C} = \overbrace{\beta_C \phi_{VC} C V_C}^{\text{lysis}} - \overbrace{m_{VC} V_C}^{\text{decay}} \tag{S6}$$

$$\text{Viruses of E: } \dot{V_E} = \overbrace{\beta_E \phi_{VE} E V_E}^{\text{lysis}} - \overbrace{m_{VE} V_E}^{\text{decay}} \tag{S7}$$

$$\begin{aligned}\text{Organic N: } \dot{x}_{on} = & -\overbrace{\frac{q_H}{\epsilon_H} \frac{\mu_H H x_{on}}{x_{on}+K_{on}}}^{\text{H growth}} + \overbrace{q_V m_{VH} V_H + q_V m_{VC} V_C + q_V m_{VE} V_E}^{\text{viral decay}} \\ & + \overbrace{(q_H - q_V \beta_H) \phi_{VH} H V_H}^{\text{H lysis by viruses}} + \overbrace{(q_C - q_V \beta_C) \phi_{VC} C V_C}^{\text{C lysis by viruses}} + \overbrace{(q_E - q_V \beta_E) \phi_{VE} E V_E}^{\text{E lysis by viruses}} \\ & + \overbrace{q_H m_{on,H} H}^{\text{loss of H}} + \overbrace{q_C m_{on,C} C}^{\text{loss of E}} + \overbrace{q_E m_{on,E} E}^{\text{loss of C}} \\ & + \overbrace{p_{on} q_H \psi_{ZH} H Z}^{\text{H grazing by Z}} + \overbrace{p_{on} q_C \psi_{ZC} C Z}^{\text{C grazing by Z}} + \overbrace{p_{on} q_E \psi_{ZE} E Z}^{\text{E grazing by Z}} \end{aligned} \tag{S8}$$

$$\begin{aligned}\text{Inorganic N: } \dot{x}_{in} = & \overbrace{-\omega(x_{in} - x_{sub})}^{\text{import}} + \overbrace{\frac{q_H(1-\epsilon_H)}{\epsilon_H} \frac{\mu_H H x_{on}}{x_{on}+K_{on}}}^{\text{H growth}} - \overbrace{\frac{q_C \mu_C C x_{in}}{x_{in}+K_{in,C}}}^{\text{C growth}} - \overbrace{\frac{q_E \mu_E E x_{in}}{x_{in}+K_{in,E}}}^{\text{E growth}} + \overbrace{q_Z m_z Z}^{\text{respiration}} \\ & + \overbrace{q_H m_{in,H} H}^{\text{respiration of H}} + \overbrace{q_C m_{in,C} C}^{\text{respiration of E}} + \overbrace{q_E m_{in,E} E}^{\text{respiration of C}} \\ & + \overbrace{p_{in} q_H \psi_{ZH} H Z}^{\text{H grazing by Z}} + \overbrace{p_{in} q_C \psi_{ZC} C Z}^{\text{C grazing by Z}} + \overbrace{p_{in} q_E \psi_{ZE} E Z}^{\text{E grazing by Z}} \end{aligned} \tag{S9}$$

$$\text{Export : } J_{out} = \overbrace{p_{ex} q_Z m_{ZP} Z^2}^{\text{consumption}} \tag{S10}$$

1.2 Model without viruses

We propose an alternative model without viruses:

$$\text{Heterotrophs: } \dot{H} = \overbrace{\frac{\mu_H H x_{on}}{x_{on} + K_{on}}}^{\text{H growth}} - \overbrace{\psi_{ZH} H Z}^{\text{grazing}} - \overbrace{m_{on,H} H}^{\text{organic loss}} - \overbrace{m_{in,H} H}^{\text{respiration}} \quad \text{(S11)}$$

$$\text{Cyanobacteria: } \dot{C} = \overbrace{\frac{\mu_C C x_{in}}{x_{in} + K_{in,C}}}^{\text{C growth}} - \overbrace{\psi_{ZC} C Z}^{\text{grazing}} - \overbrace{m_{on,C} C}^{\text{organic loss}} - \overbrace{m_{in,C} C}^{\text{respiration}} \quad \text{(S12)}$$

$$\text{Euk. autos: } \dot{E} = \overbrace{\frac{\mu_E E x_{in}}{x_{in} + K_{in,E}}}^{\text{E growth}} - \overbrace{\psi_{ZE} E Z}^{\text{grazing}} - \overbrace{m_{on,E} E}^{\text{organic loss}} - \overbrace{m_{in,E} E}^{\text{respiration}} \quad \text{(S13)}$$

$$\text{Zooplankton } \dot{Z} = \overbrace{p_g \left(\frac{q_H}{q_Z} \psi_{ZH} H Z + \frac{q_C}{q_Z} \psi_{ZC} C Z + \frac{q_E}{q_Z} \psi_{ZE} E Z \right)}^{\text{grazing}} - \overbrace{m_Z Z}^{\text{respiration}} - \overbrace{m_{ZP} Z^2}^{\text{consumption}} \quad \text{(S14)}$$

$$\text{Organic N: } \dot{x}_{on} = -\overbrace{\frac{q_H}{\epsilon_H} \frac{\mu_H H x_{on}}{x_{on} + K_{on}}}^{\text{H growth}} + \overbrace{q_H m_{on,H} H}^{\text{loss of H}} + \overbrace{q_C m_{on,C} C}^{\text{loss of E}} + \overbrace{q_E m_{on,E} E}^{\text{loss of C}}$$
$$+ \overbrace{p_{on} q_H \psi_{ZH} H Z}^{\text{H grazing by Z}} + \overbrace{p_{on} q_C \psi_{ZC} C Z}^{\text{C grazing by Z}} + \overbrace{p_{on} q_E \psi_{ZE} E Z}^{\text{E grazing by Z}} \quad \text{(S15)}$$

$$\text{Inorganic N: } \dot{x}_{in} = \overbrace{-\omega(x_{in} - x_{sub})}^{\text{import}} + \overbrace{\frac{q_H(1-\epsilon_H)}{\epsilon_H} \frac{\mu_H H x_{on}}{x_{on} + K_{on}}}^{\text{H growth}} - \overbrace{\frac{q_C \mu_C C x_{in}}{x_{in} + K_{in,C}}}^{\text{C growth}} - \overbrace{\frac{q_E \mu_E E x_{in}}{x_{in} + K_{in,E}}}^{\text{E growth}} + \overbrace{q_Z m_z Z}^{\text{respiration}}$$
$$+ \overbrace{p_{in} q_H \psi_{ZH} H Z}^{\text{H grazing by Z}} + \overbrace{p_{in} q_C \psi_{ZC} C Z}^{\text{C grazing by Z}} + \overbrace{p_{in} q_E \psi_{ZE} E Z}^{\text{E grazing by Z}}$$
$$+ \overbrace{q_H m_{in,H} H}^{\text{respiration of H}} + \overbrace{q_C m_{in,C} C}^{\text{respiration of E}} + \overbrace{q_E m_{in,E} E}^{\text{respiration of C}} \quad \text{(S16)}$$

$$\text{Export : } J_{out} = \overbrace{p_{ex} q_Z m_{ZP} Z^2}^{\text{consumption}} \quad \text{(S17)}$$

1.3 Simulation framework

All simulations are conducted in MATLAB. The two models, with and without viruses, are simulated using the same codebase. The code accepts a set of interactions and then builds a coupled system of ODE-s out of this set. Numerical integration of the coupled system of ODE-s is done using a higher order Runga-Kutta integration scheme (ode45 in MATLAB). A complete README is available and the source code is available for use without restriction as a Supplementary File and is available for download at GitHub.

2 Parameter estimation

2.1 Rationale for upper and lower bounds of model parameters

Initial parameter ranges in Table S2 were based on established values in models of aquatic food webs with viruses [27, 17] and without viruses [28, 22]. Because the current model differs in structure, some of the parameter variation may have been due to virus effects and not explicitly modeled. As such, we tried, when possible, to select broader potential parameter regimes for model selection, and utilized our targeted parameter search to identify those combinations of parameters that were compatible with biologically plausible concentrations. Below are details on the rationale for parameter range selection, organized by event type.

H growth : Heterotrophs can double on rates that vary from hourly to daily. However, as a functional group, we utilized conventions of μ_H of approximately 0.5 to 2 days^{-1}, consistent estimates ranging from $0.1 - 2$ days^{-1} [7]. The half-saturation constant, K_{on}, are based on estimates of 0.5 μM N cited in [4]. The growth efficiency ϵ_H is based on ranges established in [6] for ocean systems that have the bulk of their variation between 0.1-0.3, with biological variation outside of that range.

C growth : Cyanobacteria doubling times are approximately 12 hrs-24 hrs [29] - which we expand to include rates that correspond to approximately 0.5 to 2 days^{-1}. The half-saturation rates are chosen from the low end of reported phytoplankton ranges, which have been reported to be less than 0.2 μmol/L for natural oceanic communities, to include the range 0.05 to 1 μmol/L [9, 8].

E growth : Eukaryotic autotrophs have maximum growth rates of hours to days, with substantial variation given variation in nutrient availability. Yet, they can also have higher half-saturation constants than do cyanobacteria with estimates of $K_{in,E}$ reaching 10 μmol/L [14, 21, 15]. In practice, we select half-saturation constants in the range of reported values from 0.5 to 10μmol/L [9, 21], and maximum growth rates from 0.2 to 2 days^{-1}.

Viral lysis : Maximum values of diffusion-limited adsorption rates are approximately 2×10^{-9} L/(viruses-days) given a microbial host with 2 μm diameter and a virus with a 50 nm diameter capsid. This rate increases proportionately with host diameter and the inverse of virus effective radius. However, realized diffusion-limited adsorption rates are less for biophysical reasons, e.g., receptors are not uniformly distributed on cell surfaces, and ecological reasons, e.g., not all host cells are targeted by all viruses [18]. Here we set the upper limit of adsorption to be 5 times below the maximum. In practice, we selected a lower range two or three orders of magnitude below. Next, burst sizes vary significantly for viruses of marine microbes. We used baseline estimates that burst sizes range from 20-40 for lysis of marine heterotrophs [20]. We permitted higher upper ranges (up to 100) for infections of cyanobacteria even higher upper ranges (up to 500) for infections of microeukaryotes [3]. Upper limits of lysis in certain infections of large phytoplankton have been reported but were not included in the ranges [3].

Viral decay : Residence times of virus particles in the upper surface ocean have been estimated to be on the order of 1 day [25]. Here we permit potentially faster clearance, at a rate up to 5 days^{-1}. Intrinsic decay rates have been estimated to be on the order of weeks at room temperature [5] which sets our lower limit of decay of 0.05 days^{-1}.

Zooplankton grazing Stock and Dunne [22] implemented a Type-II functional response for grazing. Here, we utilize a Type-I functional response. These can be compared in the limit of small prey densities. Stock and Dunne [22] report:

$$\frac{I_{max,SZ}}{K_{I,SZ}} = \frac{5}{[0.68 - 1.18]} \frac{m^3}{day\, mmoles N} = 4.2 - 7.4 \frac{L}{day\, \mu mol\, N}$$

The range used here is inclusive with:

$$\frac{\psi}{q_Z} = \frac{[10^{-6} - 10^{-4}]}{[5 \times 10^{-5} 4 \times 10^{-4}]} \frac{L}{day\, \mu mol\, N} = 0.025 - 20 \frac{L}{day\, \mu mol\, N}$$

consistent with observations in [12]. The use of an extended lower limit enables the exploration of a dynamic range in which zooplankton are relatively less efficient than viruses in clearing hosts. The value of the half-saturation constant, $K_{I,SZ}$ in [22] was calibrated to their data and does not have an analogous parameter in the current model. The fixed value of 40% assimilation efficiency p_g is based on measurements of approximately 20%-40% efficiency for nano-microflagellates, with additional variation outside of this 75% confidence interval [23]. The breakdown between egestion, p_{on}, and respiration, $p_i n$, are based on outcomes reported in [19].

Grazer respiration Stock and Dunne [22] used ranges: 0.05-0.4 day^{-1} for small zooplankton and 0.02-0.16 day^{-1} for large zooplankton. Here we utilize the range $0.025 - 0.1 day^{-1}$.

Consumption by higher predators The parameter m_{ZP} enables the closure of the model, as it corresponds to consumption of the "top" predator in the current model by other, larger predators. A fraction, p_{ex} of this is transferred out of the system to higher trophic levels. The limits were calibrated to ensure model coexistence in a plausible regime.

Import The concentration of inorganic nutrient below the nutricline determined the value of x_{sub} [28]. The stable mixed layer has a turnover time ranging from 50 to 200 days. This is a feature of the assumption of a highly retentive surface microbial ecosystem as modeled here.

Nutrient levels Cyanobacteria nitrogen content can vary with strain and growth conditions, e.g., from approximately 10 to 50 fg per cell with the strains *Prochlorococcus* MED4, *Synechococcus* WH8012 and *Synechococcus* WH8013 respectively [2]. This range corresponds to 0.7×10^{-9} to 3.5×10^{-9} μmol N per cell, which is used as the basis for the selected range of 0.5-4×10^{-9}. Note that although heterotrophic cell sizes can be smaller, e.g., on the order of 20 fg per cell [7], they also differ in the nutrient content, with larger-than-expected nitrogen-to-carbon ratios [24]. Here, an identical range for nitrogen content is used for heterotrophs. Eukaryotic autotrophs spanning a size range $2-5$ μm in size are approximately 5 times large in size. This increase in physical dimension translates into an approximately two-order of magnitude increase in carbon and nitrogen content, such that we assume $q_E \approx 2\times10^{-7}$ μmol N/cell and include variation around this baseline. The nano-/micro-zooplankton group are assumed to be 1 order of magnitude larger in physical dimension, i.e., between $10-50$ μ m in size, such that they have a baseline nitrogen content of $q_Z \approx 10^{-4}$ μmol N/cell with variation around that value. Finally, viral nitrogen content, q_V, has been estimated via a first principles biophysical theory to range between $1.4\times10^{-12} - 14\times10^{-12}$ μmol N/(virus particle) for viruses whose capsids have diameters between 30-70 nm [13]. Estimates for nutrient content of viruses are based on estimates focusing on the heads of bacteriophage particles containing dsDNA without lipids. We expand this range to account for potential biological variation, in both size and structure, across all three types, ranging from 0.5×10^{-12} to 20×10^{-12} μmol N/particle.

Cellular loss We include a linear loss of phytoplankton to inorganic material to represent basal metabolic losses and a similar loss to organic material to represent exudation. Basal respiration ranges were informed by baseline estimates of 10% relative to production rates, using 1 per day production as a baseline [10] (while noting that variation can be substantial, particular during starvation periods). Basal exudation rates were informed by [1], which estimates that approximately 13% of total fixed organic matter is released. We set the upper range to be approximately 1/10 of maximum growth rates when growing at a rate of 1 per day. The model analysis is robust to choices in these ranges, so long as dominant forms of mortality are due to interactions rather than death due to loss of viability in the absence of nutrients (i.e., exudation and respiration).

2.2 Constraints on parameter combinations

Parameters within bounds, delineated above, cannot always be chosen independently. In particular, the nitrogen content of virions released upon lysis cannot exceed the nitrogen content of uninfected hosts. Hence, we find the following self-consistency conditions:

$$q_H > q_V\beta_{VH} \qquad q_C > q_V\beta_{VC} \qquad q_E > q_V\beta_{VE}$$

where β denotes the burst size of a particular viral-host interaction (denoted in the subscript). Similarly, for self-consistency, the nitrogen content of zooplankton produced due to grazing cannot exceed the nitrogen content of hosts consumed. This is ensured by using ratios of the species q-values multiplied by a fractional use. Note that the ultimate destination of organic material grazed by zooplankton must satisfy:

$$\overbrace{p_g}^{\text{growth}} + \overbrace{p_{on}}^{\text{egestion}} + \overbrace{p_{in}}^{\text{respiration}} = 1$$

Finally, at steady state export must balance import, which reflects the balance of net import of inorganic N with the export of organic N (either up the food chain or exported out of the surface layer).

3 Analytical solutions of the coexistence steady state in models with and without viruses

3.1 Algebraic solutions of the steady state for the model with viruses

The derivation of steady states is the multi-trophic model with viruses is straightforward, though not all of the algebraic solutions yield significant insight. For completeness, we present the derivation in its entirety. To begin, set equations (5-7) to 0, yielding:

$$H^* = \frac{m_{VH}}{\beta_H \phi_{VH}} \tag{S18}$$

$$C^* = \frac{m_{VC}}{\beta_C \phi_{VC}} \tag{S19}$$

$$E^* = \frac{m_{VE}}{\beta_E \phi_{VE}} \tag{S20}$$

where the asterisks denote steady state densities. Further, note that grazer dynamics only depend on Z and the densities of the three microbial guilds. Hence, Z^* can also be solved:

$$Z^* = \frac{\frac{p_g}{q_Z}\left(q_H \psi_{ZH} H^* + q_C \psi_{ZC} C^* + q_E \psi_{ZE} E^*\right) - m_Z}{m_{ZP}} \tag{S21}$$

Because of nutrient balance, we also have a solution for x_{in}^*:

$$x_{in}^* = x_{sub} - \frac{q_Z m_{ZP} \left(Z^*\right)^2}{\omega} \tag{S22}$$

With this in hand, we also have solutions for the viral levels of cyanobacteria and eukaryotes:

$$V_C^* = \frac{\frac{\mu_C x_{in}^*}{x_{in}^* + K_{in,C}} - \psi_{ZC} Z^* - m_{in,C} - m_{on,C}}{\phi_{VC}} \tag{S23}$$

$$V_E^* = \frac{\frac{\mu_E x_{in}^*}{x_{in}^* + K_{in,E}} - \psi_{ZE} Z^* - m_{in,E} - m_{on,E}}{\phi_{VE}} \tag{S24}$$

Finally, there are two equations that we can use to help solve for x_{on}^* and V_H^*. First, from the dynamics of heterotrophs (Equation 1):

$$\frac{\mu_H x_{on}^*}{x_{on}^* + K_{on}} = \phi_{VH} V_H^* + \psi_{ZH} Z^* + m_{in,H} + m_{on,H} \tag{S25}$$

This implies that:

$$x_{on}^* = \frac{K_{on}\left(\phi_{VH} V_H^* + \psi_{ZH} Z^* + m_{in,H} + m_{on,H}\right)}{\mu_H - \phi_{VH} V_H^* - \psi_{ZH} Z^* - m_{in,H} - m_{on,H}} \tag{S26}$$

This equation is implicit in V_H^*, which we solve next. Then, recalling the dynamics of organic N:

$$\begin{aligned}
\text{Organic N: } \dot{x}_{on} = & -\overbrace{\frac{q_H}{\epsilon_H}\frac{\mu_H H x_{on}}{x_{on}+K_{on}}}^{\text{H growth}} + \overbrace{q_V m_{VH} V_H + q_V m_{VC} V_C + q_V m_{VE} V_E}^{\text{viral decay}} \\
& + \overbrace{(q_H - q_V \beta_H)\phi_{VH} H V_H}^{\text{H lysis by viruses}} + \overbrace{(q_C - q_V \beta_C)\phi_{VC} C V_C}^{\text{C lysis by viruses}} + \overbrace{(q_E - q_V \beta_E)\phi_{VE} E V_E}^{\text{E lysis by viruses}} \\
& + \overbrace{q_H m_{on,H} H}^{\text{loss of H}} + \overbrace{q_C m_{on,C} C}^{\text{loss of E}} + \overbrace{q_E m_{on,E} E}^{\text{loss of C}} \\
& + \overbrace{p_{on} q_H \psi_{ZH} H Z}^{\text{H grazing by Z}} + \overbrace{p_{on} q_C \psi_{ZC} C Z}^{\text{C grazing by Z}} + \overbrace{p_{on} q_E \psi_{ZE} E Z}^{\text{E grazing by Z}}
\end{aligned} \tag{S27}$$

we replace the first term and isolate V_H^* (ignoring the * for now, but we will re-insert them later):

$$\begin{aligned} V_H\left(\frac{q_H\phi_{VH}H}{\epsilon_H} - q_V m_{VH} - (q_H - q_V\beta_H)\phi_{VH}H\right) &= -\frac{q_H H}{\epsilon_H}\left(\psi_{ZH}Z + m_{in,H} + m_{on,H}\right) + q_V m_{VC}V_C + q_V m_{VE}V_E \\ &\quad +(q_C - q_V\beta_C)\phi_{VC}CV_C + (q_E - q_V\beta_E)\phi_{VE}EV_E \\ &\quad +p_{on}q_H\psi_{ZH}HZ + p_{on}q_C\psi_{ZC}CZ + p_{on}q_E\psi_{ZE}EZ \\ &\quad +q_H m_{on,H}H + q_C m_{on,C}C + q_E m_{on,E}E \\ &\equiv \aleph \end{aligned} \tag{S28}$$

And so, omitting * on the right hand side given the size of the equation):

$$V_H^* = \frac{\aleph}{\frac{q_H\phi_{VH}H}{\epsilon_H} - q_V m_{VH} - (q_H - q_V\beta_H)\phi_{VH}H} \tag{S29}$$

Our finding of an algebraic solution facilitates rapid evaluation of the dependence of steady state densities on parameters. We do not claim that the algebraic forms, in and of themselves, are necessarily insightful.

3.2 Algebraic solutions of the virus-free steady state

As discussed in the main text, viruses act to enrich for diversity. Without viruses, then either the cyanophage or the eukaryotic phytoplankton go extinct. From a model standpoint, the reason is that both cyanophage and the eukaryotic phytoplankton functional groups compete for the same (single) resource and are subject to the same grazer. If we had used different functional responses or multiple resources then coexistence could be possible, even without viruses. Hence, let us for a moment consider the case where $C^* = 0$ and all the viruses are zero. In that case, and dropping the * for now, we recognize that the steady state solution can be found as follows.

First, because of nutrient balance, we also have a solution for x_{in}:

$$x_{in} = x_{sub} - \frac{q_Z m_{ZP} Z^2}{\omega} \tag{S30}$$

However, because of the conditions implied by $\dot{E} = 0$, then

$$Z = \frac{1}{\psi_{ZE}}\left[\frac{\mu_E x_{in}}{(x_{in} + K_{in,E})} - m_{on,E} - m_{in,E}\right] \tag{S31}$$

which can substitute back to yield (with $m_E \equiv m_{in,E} + m_{on,E}$):

$$x_{in} = x_{sub} - \frac{q_Z m_{ZP}}{\omega\psi_{ZE}^2}\left[\left(\frac{\mu_E x_{in}}{x_{in} + K_{in,E}}\right)^2 - \frac{2m_E\mu_E x_{in}}{x_{in} + K_{in,E}} + m_E^2\right].$$

Noting that $\mathbb{C} = \frac{q_Z m_{ZP}}{\omega\psi_{ZE}^2}$, we can rewrite this as:

$$x_{in} = x_{sub} - \mathbb{C}\left[\left(\frac{\mu_E x_{in}}{x_{in} + K_{in,E}}\right)^2 - \frac{2m_E\mu_E x_{in}}{x_{in} + K_{in,E}} + m_E^2\right].$$

and re-arranging terms yields:

$$x_{in}\left(x_{in} + K_{in,E}\right)^2 = x_{sub}\left(x_{in} + K_{in,E}\right)^2 - \mathbb{C}\mu_E^2 x_{in}^2 + 2\mathbb{C}m_E\mu_E x_{in}\left(x_{in} + K_{in,E}\right) - \mathbb{C}m_E^2\left(x_{in} + K_{in,E}\right)^2. \tag{S32}$$

We then expand this out in like terms such that

$$x_{in}^3 + C_2 x_{in}^2 + C_1 x_{in} + C_0 = 0$$

where

$$C_2 = 2K_{in,E} - x_{sub} + \mathbb{C}\mu_E^2 - 2\mathbb{C}m_E\mu_E + \mathbb{C}m_E^2 \tag{S33}$$

$$C_1 = K_{in,E}^2 - 2x_{sub}K_{in,E} - 2\mathbb{C}m_E\mu_E K_{in,E} + 2\mathbb{C}m_E^2 K_{in,E} \tag{S34}$$

$$C_0 = -x_{sub}K_{in,E}^2 + \mathbb{C}m_E^2 K_{in,E}^2 \tag{S35}$$

We can then solve this cubic algebraically to find x^*_{in}. It is well known that this may admit 1 or 3 real solutions (ignoring degenerate cases). Defining $\Delta = 18C_2C_1C_0 - 4C_2^3C_0 + C_2^2C_1^2 - 4C_1^3 - 27C_0^2$, we note that when $\Delta > 0$ then there will be three real roots and when $\Delta < 0$ there will be one real root. The sign of Δ and the positivity (or not) of roots is examined numerically.

Next, we work progressively to solve the remainder of the steady state values algebraically, implicit given the value of x^*_{in}. First, zooplankton:

$$Z = \left[\frac{\omega\left(x_{sub} - x_{in}\right)}{q_z m_{ZP}}\right]^{1/2} \tag{S36}$$

then using the $\dot{H} = 0$ condition:

$$x_{on} = \frac{K_{on}\left(\psi_{ZH}Z + m_{on,H} + m_{in,H}\right)}{\mu_H - \psi_{ZH}Z - m_{on,H} - m_{in,H}} \tag{S37}$$

then using the $\dot{x}_{on}$ equation and the relationship implied by the $\dot{Z}$ equation, yields:

$$\begin{aligned} H\left[\frac{q_H \mu_H x_{on}}{\epsilon_H\left(x_{on} + K_{on}\right)} - q_H m_{on,H} + q_E m_{on,E}\frac{q_H \psi_{ZH}}{q_E \psi_{ZE}}\right] &= \frac{p_{on} q_Z Z\left(m_Z + m_{ZP}Z\right)}{p_g} \\ &\quad + q_E m_{on,E}\frac{q_Z\left(m_Z + m_{ZP}Z\right)}{p_g q_E \psi_{ZE}} \end{aligned} \tag{S38}$$

such that

$$H = \frac{\frac{p_{on} q_Z Z(m_Z + m_{ZP}Z)}{p_g} + q_E m_{on,E}\frac{q_Z(m_Z + m_{ZP}Z)}{p_g q_E \psi_{ZE}}}{\left[\frac{q_H \mu_H x_{on}}{\epsilon_H(x_{on} + K_{on})} - q_H m_{on,H} + q_E m_{on,E}\frac{q_H \psi_{ZH}}{q_E \psi_{ZE}}\right]} \tag{S39}$$

and finally,

$$E = \left(m_{ZP}Z + m_Z - p_g q_H \psi_{ZH} H / q_Z\right)\frac{q_Z}{p_g q_E \psi_{ZE}} \tag{S40}$$

Due to the symmetry of the model construction, a similar steady state is found in cases where $E \to 0$ and, instead $C^* > 0$.

3.3 Stability of equilibria

We evaluated the expected stability of feasible fixed points. The stability was evaluated by calculating the Jacobian at the equilibrium. An analytic expression for the Jacobian was evaluated automatically based on the first derivatives of the predicted steady states. The Jacobian was then evaluated given parameters and predicted densities for each fixed point (complete expressions available as a source file in the software release). The stability of the steady state was classified into stable nodes, stable spirals, unstable nodes and unstable spirals. The conditions for each were based on the real and imaginary components of the largest eigenvalue. Stability was determined based on whether the largest eigenvalue was negative (stable) or positive (unstable). Nodes had no imaginary component where spirals did. For example, a fixed point was considered to be a stable spiral if the real part of the largest eigenvalue was negative, but at least one eigenvalue had a non-zero imaginary term.

4 Sensitivity analysis

4.1 Latin-hypercube sampling of parameters

A LHS approach was used to sample parameter space [16]. Of the 38 parameters, 35 were allowed to vary. The three that remained fixed for all models were related to the fractional allocation of prey biomass consumed by zooplankton, i.e., p_g, p_{on}, and p_{in}. We used uniform distributions in logarithmic space where the lower and upper bounds for all parameters are shown in Table S1. We used 10 resamples for each of 10^5 independent selection of midpoints within the stratified 35-dimensional parameter space, for a total of 10^6 random parameter sets.

4.2 Identifying baseline parameters consistent with "known" system densities

Initial parameters were chosen by first establishing approximate values for all parameters from the literature and from first-principle derivations. Given that parameters represent functional groups, these initial values were used to seed an automatic approach to finding parameter sets compatible with steady state densities commonly observed in the surface oceans. Because there are 38 parameters and multiple constraints among parameters, the following automated procedure was developed:

1. An initial guess for all parameters was chosen, denoted as x_0. This value was selected via LHS approach (see Table S2).

2. Lower and upper bounds were selected for all parameters, denoted as x_l and x_h, respectively. The bounds were in most cases a factor of 2 below and above the initial guesses, except for interaction rates which are often most uncertain, which were allowed to vary by 3 or 4 orders of magnitude.

3. A nonlinear minimization routine (fmincon in MATLAB) was utilized to find an optimal set of 38 parameters which we aggregate as a single vector $\vec{\theta}$. The objective was to find a parameter set $\vec{\theta^*}$ that satisfied:

$$\vec{\theta^*} \in \vec{\theta} \quad \text{such that} \quad \text{Min} \sum_{i=1}^{13} \log \left(\frac{y_i^t(\vec{\theta})}{y_i^*} \right)^2 \tag{S41}$$

where $y^t(\theta)$ is the model output given a candidate parameter set θ and y^* is the vector of target densities, i.e.,

$$y_d^* \equiv (H^*, C^*, E^*, Z^*, V_H^*, V_C^*, V_E^*, x_{on}^*, x_{in}^*, \ldots) \tag{S42}$$

augmented by four additional features, the ratio of virus to hosts, and the fraction of mortality of H, C and E due to viruses, for a total of 13 targets. The index i denotes one of the 13 components of the model output. The choice of this is meant to weight all output features equally, while balancing deviations on vastly different scales.

The following are the "target" set of densities utilized in this procedure:

$$\begin{aligned}
H^* &= 2 \times 10^8 \\
C^* &= 2 \times 10^8 \\
E^* &= 2 \times 10^6 \\
Z^* &= 5 \times 10^4 \\
V_H^* &= 2 \times 10^9 \\
V_C^* &= 2 \times 10^9 \\
V_E^* &= 2 \times 10^7 \\
x_{on}^* &= 5 \\
x_{in}^* &= 0.1 \\
\text{Ratio of viruses to hosts} &= 10, \\
\text{Fractional mortality due to viruses of } H &= 0.5 \\
\text{Fractional mortality due to viruses of } C &= 0.25 \\
\text{Fractional mortality due to viruses of } E &= 0.1
\end{aligned}$$

Here, the densities are in units of particles/L with the exception of x_{on}^* and x_{in}^* which are in units of μmol/L. The nonlinear optimization process yielded parameter sets, of which the top 5% are seen in Figure 4.

4.3 Sloppy-stiff analysis

Parameters in a model can have substantially different effects on model output. One approach to characterize these effects is via a form of sensitivity analysis termed sloppy-stiff analysis [11, 26]. Parameters which cause large changes to output given small variations in their values are referred to as stiff. In contrast, parameters which cause small changes to output given small variations in their values are referred to as sloppy. The information necessary to characterize the relative sloppiness and stiffness of all parameters can be derived from a Hessian matrix, where the ij-th element is defined as:

$$\mathcal{H}_{ij} = \sum_n \frac{\partial \log y_n}{\partial \log \theta_i}\bigg|_{\vec{\theta}^*} \frac{\partial \log y_n}{\partial \log \theta_j}\bigg|_{\vec{\theta}^*} = \sum_n \frac{\theta_i^* \theta_j^*}{(y_n^*)^2} \frac{\partial y_n}{\partial \theta_i}\bigg|_{\vec{\theta}^*} \frac{\partial y_n}{\partial \theta_j}\bigg|_{\vec{\theta}^*}. \quad \text{(S43)}$$

There are n total variables in the model and the reference parameters are represented by a vector $\vec{\theta}$. The bar notation emphasizes that the partial derivatives are evaluated at the reference parameter set. The argument of the sum is made unitless by a prefactor involving the reference parameter set, $\vec{\theta}^*$ and the respective equilibrium values of the variables, $\vec{y}^*$. Each element of the Hessian is interpreted as the sum of the relative changes of each variable due to relative changes of the respective parameters. Here, we considered perturbations to all parameters except p_g, p_{on}, and p_{in} due to the algebraic constraint restricting their values to a subspace. Moreover, we focus the sloppy-stiff analysis on the reference parameter set associated with the lowest deviation from the target densities.

Eigenvalues of the Hessian quantify the total amount of change to all variables given a perturbation to a combination of parameters defined by the respective eigenvector. The eigenvalue spectrum shows that perturbations to some combinations of parameters cause a majority of the effect on variables (Figure S6). Each eigenvalue corresponds to some combination of parameters. In our case, the eigenvalue spectrum is dominated by the principle eigenvalue. To get a sense of which parameters are stiff and sloppy We compute the orientation of the principle eigenvector in parameter space to identify those parameters that are most "stiff" (Figure S7). For this reference set and metric of perturbations, small values associated with the cyanobacteria processes and with zooplankton respiration are predicted to have the largest effect on output.

References

[1] S. B. Baines and M.L. Pace. The production of dissolved organic matter by phytoplankton and its importance to bacteria: patterns across marine and freshwater systems. *Limnology and Oceanography*, 36:1078–1090, 1991.

[2] S. Bertilsson, O. Berglund, D. M. Karl, and S. W. Chisholm. Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. *Limnology and Oceanography*, 48:1721–1731, 2003.

[3] Chris M. Brown, Janice E. Lawrence, and Douglas A. Campbell. Are phytoplankton population density maxima predictable through analysis of host and viral genomic DNA content ? *J. Mar. Biol. Assn. U. K.*, 86:491–498, 2006.

[4] J. R. Christian and T. R. Anderson. Modeling DOM biogeochemistry. In D. A. Hansell and C. A. Carlson, editors, *Biogeochemistry of marine dissolved organic matter*, pages 717–755, New York, 2002. Academic Press.

[5] Marianne De Paepe and François Taddei. Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS biology*, 4:e193, 2006.

[6] Paul A. del Giorgio and Jonathan J. Cole. Bacterial growth efficiency in natural aquatic systems. *Annual Review of Ecology and Systematics*, 29(1):503–541, 1998.

[7] H. W. Ducklow. Bacterial production and biomass in the oceans. In D. L. Kirchman, editor, *Microbial Ecology of the Oceans*, pages 85–120, New York, 2000. John Wiley and Sons.

[8] Kyle F. Edwards, Mridul K. Thomas, Christopher a. Klausmeier, and Elena Litchman. Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton. *Limnology and Oceanography*, 57:554–566, 2012.

[9] R. W. Eppley, J. N. Rogers, and J. J. McCarthy. Half-saturation constants for uptake of nitrate and ammonium by marine phytoplankton. *Limnol. Oceanogr.*, 14:912–920, 1969.

[10] R. J. Geider. Respiration: taxation without representation. In P. G. Falkowski, A. D. Woodhead, and K. Vivirito, editors, *Primary Productivity and Biogeochemical Cycles in the Sea*, New York, 1992. Plenum Press.

[11] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol*, 3:e189, 2007.

[12] P. J. Hansen, P. K. Bjornsen, and B. W. Hansen. Zooplankton grazing and growth: scaling within the 2-2000-mm body size range. *Limnology and Oceanography*, 42:687–704, 1997.

[13] Luis F Jover, T. Chad Effler, A. Buchan, Steven W. Wilhelm, and Joshua S. Weitz. The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nature Reviews Microbiology*, 12:519–528, 2014.

[14] C A Klausmeier, E Litchman, and S A Levin. Phytoplankton growth and stoichiometry under multiple nutrient limitation. *Limnology and Oceanography*, 49:1463–1470, 2004.

[15] Elena Litchman, Christopher A. Klausmeier, Oscar M. Schofield, and Paul G. Falkowski. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecology Letters*, 10(12):1170–1181, 2007.

[16] M. D McKay, R. J. Beckman, and W. J Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239–245, 1979.

[17] Takeshi Miki, Takefumi Nakazawa, Taichi Yokokawa, and Toshi Nagata. Functional consequences of viral impacts on bacterial communities: a food-web model analysis. *Freshwater Biology*, 53:1142–1153, June 2008.

[18] A. G. Murray and G. A. Jackson. Viral dynamics: a model of the effects of size, shape, motion and abundance of single-celled planktonic organisms and other particles. *Marine Ecology Progress Series*, 89:103–116, 1992.

[19] T Nagata. Production mechanisms of dissolved organic matter. In D.L. Kirchman, editor, *Microbial Ecology of the Oceans*, pages 121–152, New York, 2000. John Wiley and Sons, Inc.

[20] P Parada, Gerhard J Herndl, and Markus G Weinbauer. Viral burst size of heterotrophic prokaryotes in aquatic systems. *J. Mar. Biol. Ass. UK*, 86:613–621, 2006.

[21] Géraldine Sarthou, Klaas R. Timmermans, Stéphane Blain, and Paul Tréguer. Growth physiology and fate of diatoms in the ocean: a review. *Journal of Sea Research*, 53:25 – 42, 2005. Iron Resources and Oceanic Nutrients - Advancement of Global Environmental Simulations.

[22] C Stock and J Dunne. Controls on the ratio of mesozooplankton production to primary production in marine ecosystems. *Deep Sea Research I*, 57:95 – 112, 2010.

[23] D. Straile. Gross growth efficiencies of protozoan and metazoan zooplankton and their dependence on food concentration, predator-prey weight ratio, and taxonomic group. *Limnology and Oceanography*, 42:1375–1385, 1997.

[24] Curtis A. Suttle. Marine viruses - major players in the global ecosystem. *Nature Reviews Microbiology*, 5:801–812, OCT 2007.

[25] Curtis A Suttle and Feng Chen. Mechanisms and rates of decay of marine viruses in seawater. *Appl. Env. Micro.*, 58:3721–3729, 1992.

[26] Bradford Taylor, Tae J. Lee, and Joshua S. Weitz. A guide to sensitivity analysis of quantitative models of gene expression dynamics. *Methods*, 62(1):109 – 120, 2013.

[27] T. F. Thingstad. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography*, 45:1320–1328, 2000.

[28] Richard G. Williams and Michael J. Follows. *Ocean Dynamics and the Carbon Cycle Principles and Mechanisms*. Cambridge University Press, Cambridge, UK, 2011.

[29] Erik R. Zinser, Debbie Lindell, Zackary I. Johnson, Matthias E. Futschik, Claudia Steglich, Maureen L. Coleman, Matthew A. Wright, Trent Rector, Robert Steen, Nathan McNulty, Luke R. Thompson, and Sallie W. Chisholm. Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph, *Prochlorococcus*. *PLoS ONE*, 4(4):e5135, 04 2009.

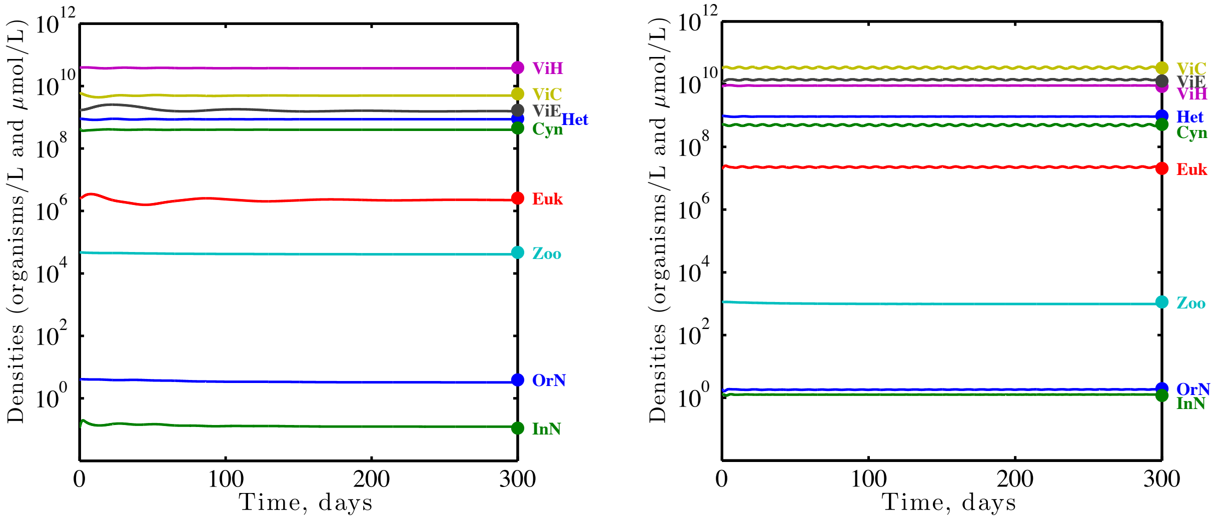

Figure S1: Dynamics of the model near putative steady states. (A) Dynamics shown to asymptotically converge to steady state. (B) Dynamics shown to diverge from a steady state leading to a limit cycle. In both cases, filled circles denoted at 300 days denote predicted equilibrium values of populations and nutrients, given parameters listed in Supplementary Table S1.

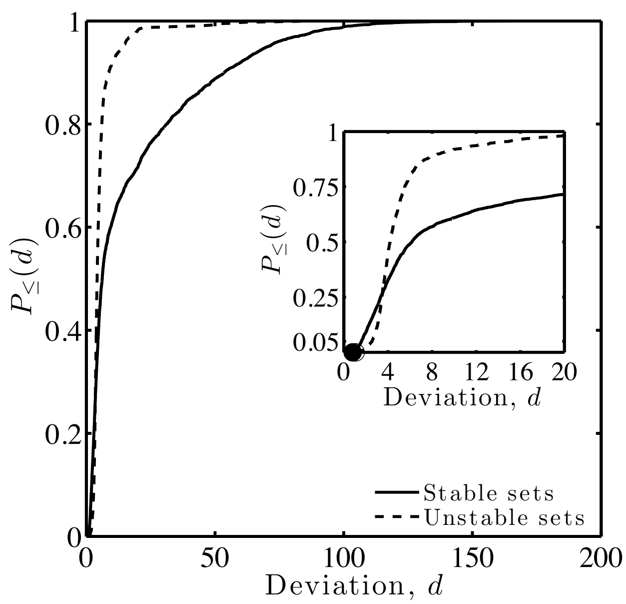

Figure S2: Cumulative distributions, $p_{\leq}(d)$, of deviations d from the target set of 13 densities and indices associated with equilibria of the multi-trophic ecosystem model (see text for main details). Distributions are associated with fixed points that are stable (solid line) or unstable (dashed line). The inset focuses on deviations $d \leq 20$, where it is evident that smaller deviations are associated with fixed points that are stable. 94% of the top 5% of parameters sets corresponded to equilibrium that were locally stable. Further, we find that the distribution of deviations associated with stable and unstable fixed points are significantly unequal ($p \ll 10^{-10}$). In particular, as is evident in Figure S2, the cumulative distribution of the deviations, d, has significantly greater mass at low values of d for equilibria associated with stable fixed points than it does for those associated with unstable fixed points.

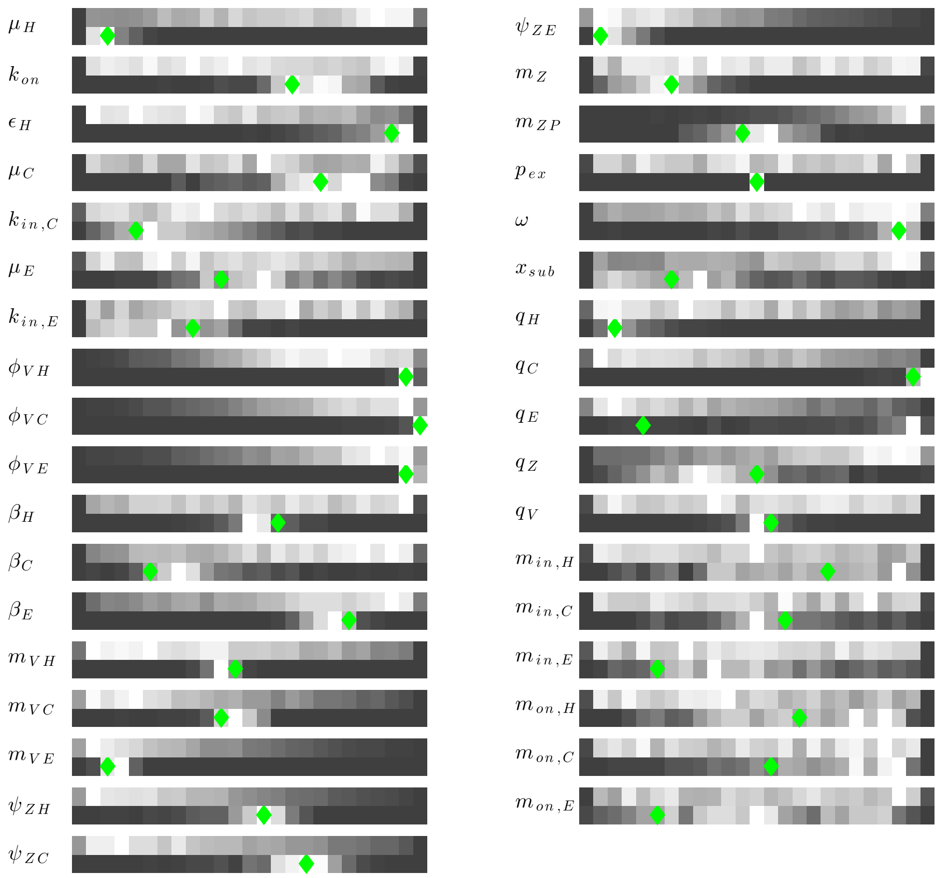

Figure S3: Distributions of parameter values obtained via nonlinear optimization. The distribution correspond to the top 5% of model feature output. The barcodes represent intensity histograms (white denotes greater number of replicates). There are two “barcodes” associated with each parameter. The top barcode corresponds to the distribution of parameters associated with feasible parameter sets obtained via the LHS sampling. The bottom barcode corresponds to the distribution of parameters in the targeted parameter sets obtained via nonlinear optimization. The green diamond denotes the value of the parameter associated with the best hit model replicate.

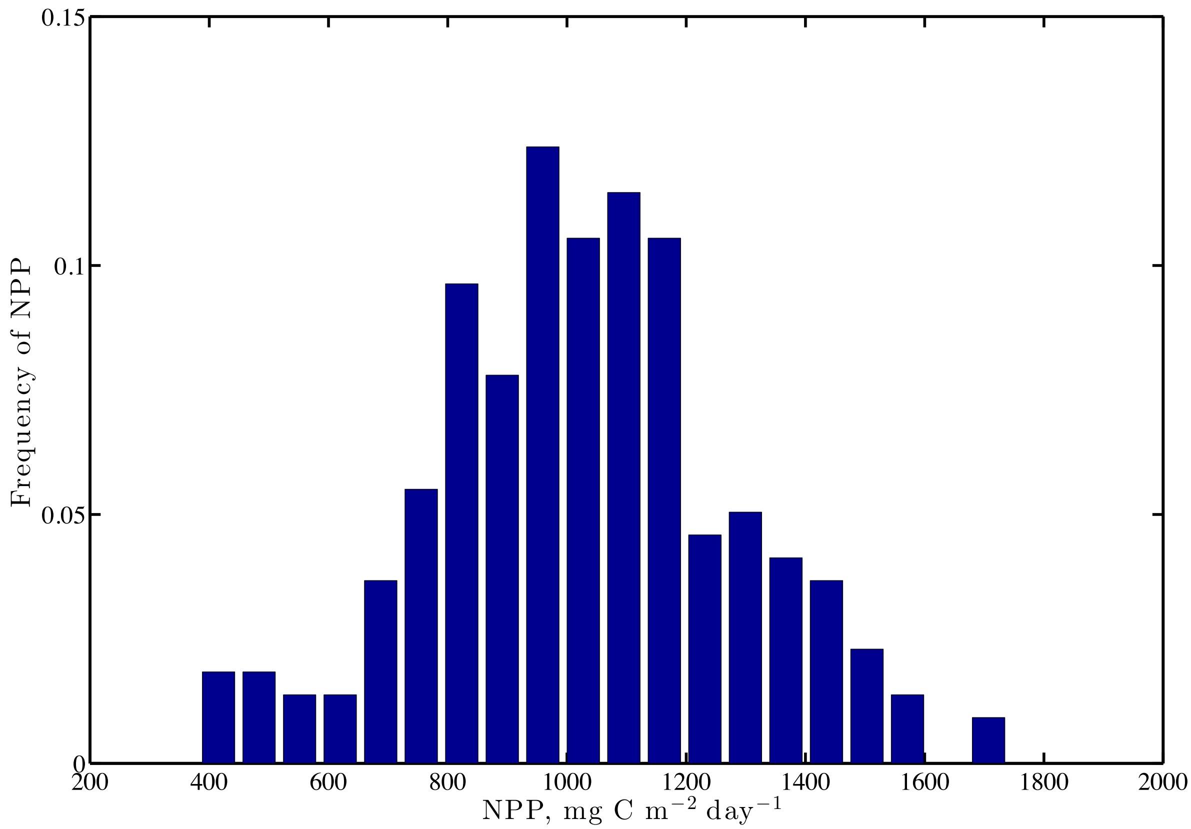

Figure S4: NPP for the top 5% of model replicates. NPP in units of mg C m^{-2} days^{-1} was calculated by multiplying the total net primary productivity in units of μmol N L^{-1} days^{-1} by the conversion factors (106/16) (denoting a baseline C:N ratio), 25 m (denoting a euphotic zone, above the nutricline), and 12.01 (denoting g C per mole). Note that moving from μmol to mmol and L to m^3 involving canceling factors of 1000. The NPP has a mean of 1030 with a standard deviation of 260.

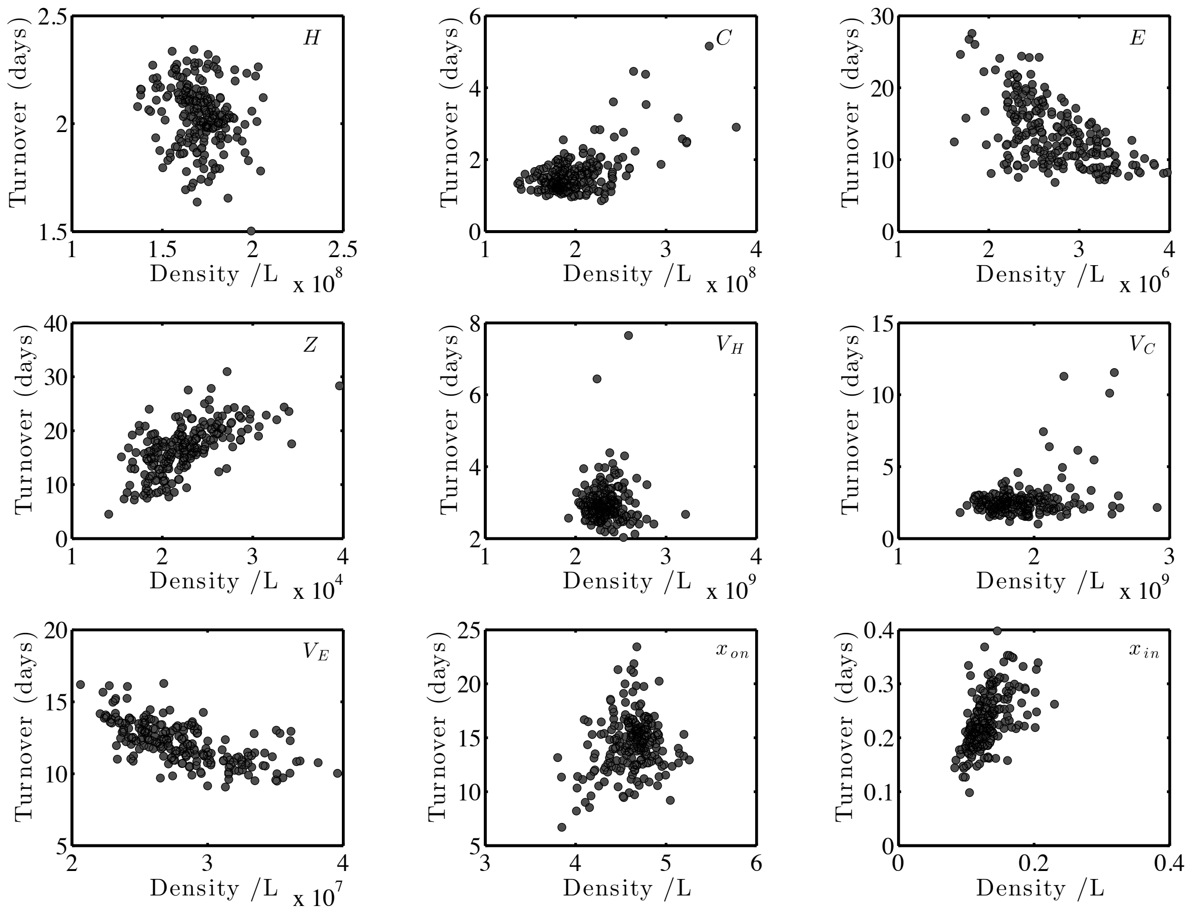

Figure S5: Turnover time, in days, for all dynamic variables, as estimated for the targeted parameter sets. In general, given a nonlinear dynamical system of the form $\dot{x} = a - kx$ where a is an input rate and k is a turnover rate, the expectation is that the variable x has a residence time (or equivalently, turnover time) of $\tau = 1/k$. Note that the equilibrium, $x^* = a/k$. Hence, the turnover time can be estimated based on the loss rate $\tau_- = \frac{x^*}{kx^*} = \frac{1}{k}$ or the input rate $\tau_+ = \frac{x^*}{a} = \frac{a}{ak} = \frac{1}{k}$. These must be equal at equilibrium as inputs balance losses. We apply this same principle to estimate the turnover time for each of the nine variables in the model. In each case, the dynamical system in Eqs. (A1-A9) can be separated into input terms, I, and loss terms, D. Then, for each equilibrium $x^*(\theta)$ associated with a parameter set θ, the turnover time was estimated at the equilibrium.

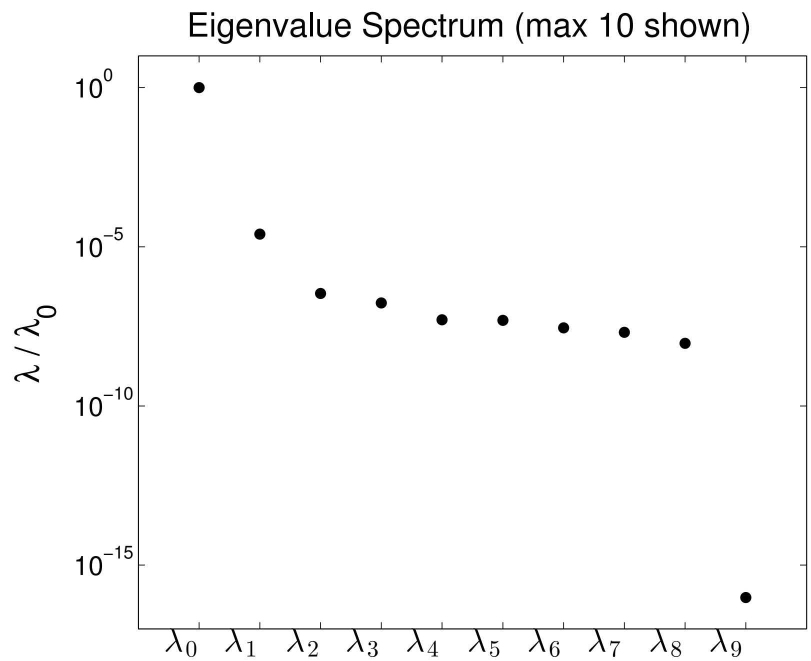

Figure S6: Eigenvalue Spectrum of parametric Hessian matrix. Values are normalized relative to the principle eigenvalue. Only the first 10 eigenvalues of the spectrum are shown. Note that the most of the perturbations in model output are due to deviations along a single, principle eigenvector, associated with the principle eigenvalue. Analysis done by considering perturbations to all parameters except p_g, p_{on}, and p_{in}.

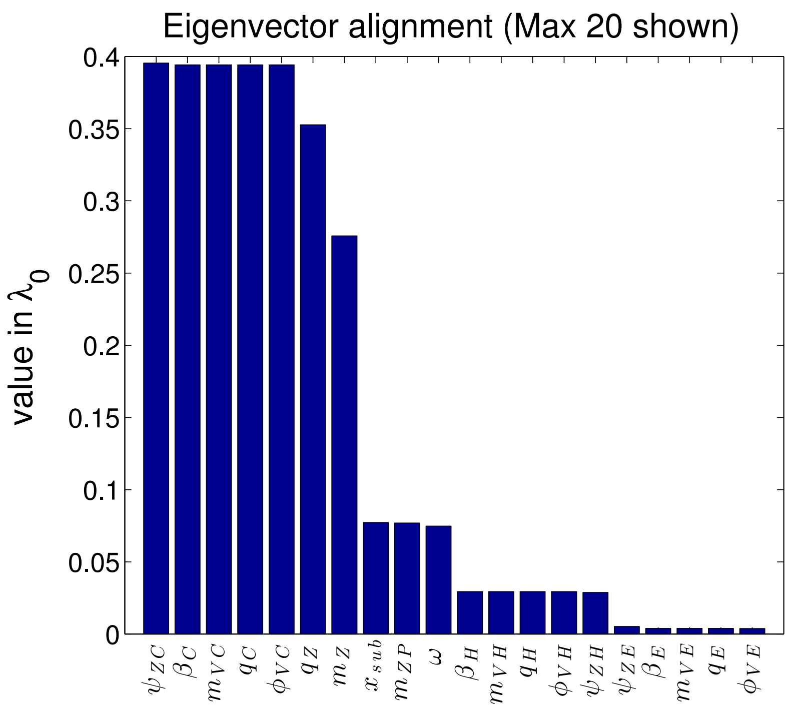

Figure S7: Projection of biological parameters along the principle eigenvector. The higher values correspond to stiffer parameters. Only the 20 stiffest parameters are shown. Analysis done by considering perturbations to all parameters except p_g, p_{on}, and p_{in}.

Event	Variable	Meaning	Units	Attracting	Limit cycle
H growth	μ_H	Max H growth rate	day^{-1}	0.51	0.61
	K_{on}	Half-saturation constant	μmol/L	0.94	0.68
	ϵ_H	Efficiency	N/A	0.20	0.19
C growth	μ_C	Max C growth rate	day^{-1}	1.97	0.68
	$K_{in,C}$	Half-saturation constant	μmol/L	0.053	0.059
E growth	μ_E	Max E growth rate	day^{-1}	7.5	5.4
	$K_{in,E}$	Half-saturation constant	μmol/L	5.4	5.8
Viral lysis	ϕ_{VH}	Lysis rate	L/(virus·day)	8.0×10^{-12}	4.3×10^{-11}
	ϕ_{VC}	Lysis rate	ibid	9.9×10^{-11}	1.9×10^{-11}
	ϕ_{VE}	Lysis rate	ibid	5.8×10^{-11}	6.7×10^{-11}
	β_H	Burst size	N/A	25	15
	β_C	Burst size	N/A	16	82
	β_E	Burst size	N/A	440	370
Viral decay	m_{VH}	Decay rate	day^{-1}	0.17	0.59
	m_{VC}	Decay rate	day^{-1}	0.63	0.76
	m_{VE}	Decay rate	day^{-1}	0.058	0.56
Zooplankton grazing	ψ_{ZH}	Grazing rate	L/(zoopl·day)	1.9×10^{-6}	2.2×10^{-5}
	ψ_{ZC}	Grazing rate	ibid	2.1×10^{-5}	4.2×10^{-5}
	ψ_{ZE}	Grazing rate	ibid	1.5×10^{-6}	2.8×10^{-5}
	p_g	Fraction for growth	N/A	0.4	0.4
	p_{on}	Fraction egested	N/A	0.3	0.3
	p_{in}	Fraction respired	N/A	0.3	0.3
Grazer respiration	m_Z	Basal respiration	day^{-1}	0.048	0.051
Consumption by higher predators	m_{ZP}	Mortality rate	L/(cells·day)	3.3×10^{-7}	2.8×10^{-5}
	p_{ex}	Fraction exported	N/A	0.49	0.37
Import	ω	Surface-deep mixing rate	1/day	0.016	0.077
	x_{sub}	Deep inorganic N conc.	μmol N/L	7.7	2.67
Nutrient levels	q_H	Nitrogen content of H	μmol N/cell	8.8×10^{-10}	1.4×10^{-9}
	q_C	Nitrogen content of C	ibid	3.9×10^{-9}	3.0×10^{-9}
	q_E	Nitrogen content of E	ibid	2.1×10^{-7}	6.7×10^{-8}
	q_Z	Nitrogen content of Z	ibid	2.3×10^{-4}	3.9×10^{-4}
	q_V	Nitrogen content of V	ibid	2.0×10^{-12}	2.2×10^{-12}
Cellular loss	$m_{in,H}$	H respiration	day^{-1}	0.0069	0.0067
	$m_{in,C}$	C respiration	day^{-1}	0.0050	0.016
	$m_{in,E}$	E respiration	day^{-1}	0.0043	0.0023
	$m_{on,H}$	H organic loss	day^{-1}	0.014	0.034
	$m_{on,C}$	C organic loss	day^{-1}	0.022	0.0063
	$m_{on,E}$	E organic loss	day^{-1}	0.011	0.022

Table S1: Parameters used in simulations of Figure S1A (attracting) and Figure S1B (limit cycle) ecosystem dynamics.

Event	Variable	Meaning	Units	Lower Bound	Upper bound
H growth	μ_H	Max H growth rate	day^{-1}	0.50	2
	K_{on}	Half-saturation constant	μmol/L	0.25	1
	ϵ_H	Efficiency	N/A	0.05	0.2
C growth	μ_C	Max C growth rate	day^{-1}	0.5	2
	$K_{in,C}$	Half-saturation constant	μmol/L	0.05	1
E growth	μ_E	Max E growth rate	day^{-1}	0.2	2
	$K_{in,E}$	Half-saturation constant	μmol/L	0.5	10
Viral lysis	ϕ_{VH}	Lysis rate	L/(virus·day)	10^{-13}	10^{-10}
	ϕ_{VC}	Lysis rate	ibid	10^{-13}	10^{-10}
	ϕ_{VE}	Lysis rate	ibid	10^{-12}	10^{-10}
	β_H	Burst size	N/A	12.5	50
	β_C	Burst size	N/A	12.5	100
	β_E	Burst size	N/A	125	500
Viral decay	m_{VH}	Decay rate	day^{-1}	0.05	5
	m_{VC}	Decay rate	day^{-1}	0.05	5
	m_{VE}	Decay rate	day^{-1}	0.05	5
Zooplankton grazing	ψ_{ZH}	Grazing rate	L/(zoopl·day)	10^{-6}	10^{-4}
	ψ_{ZC}	Grazing rate	ibid	10^{-6}	10^{-4}
	ψ_{ZE}	Grazing rate	ibid	10^{-6}	10^{-4}
	p_g	Fraction for growth	N/A	0.4	0.4
	p_{on}	Fraction egested	N/A	0.3	0.3
	p_{in}	Fraction respired	N/A	0.3	0.3
Grazer respiration	m_Z	Basal respiration	day^{-1}	0.025	0.1
Consumption by higher predators	m_{ZP}	Mortality rate	L/(cells·day)	10^{-8}	10^{-4}
	p_{ex}	Fraction exported	N/A	0.25	1
Import	ω	Surface-deep mixing rate	1/day	0.005	0.02
	x_{sub}	Deep inorganic N conc.	μmol N/L	2.5	10
Nutrient levels	q_H	Nitrogen content of H	μmol N/cell	5×10^{-10}	4×10^{-9}
	q_C	Nitrogen content of C	ibid	5×10^{-10}	4×10^{-9}
	q_E	Nitrogen content of E	ibid	5×10^{-8}	4×10^{-7}
	q_Z	Nitrogen content of Z	ibid	5×10^{-5}	4×10^{-4}
	q_V	Nitrogen content of V	ibid	0.5×10^{-12}	20×10^{-12}
Cellular loss	$m_{in,H}$	H respiration	day^{-1}	0.001	0.1
	$m_{in,C}$	C respiration	day^{-1}	0.001	0.1
	$m_{in,E}$	E respiration	day^{-1}	0.001	0.1
	$m_{on,H}$	H organic loss	day^{-1}	0.005	0.1
	$m_{on,C}$	C organic loss	day^{-1}	0.005	0.1
	$m_{on,E}$	E organic loss	day^{-1}	0.005	0.1

Table S2: Parameter ranges used for LHS sampling of initial parameters.