

RESEARCH ARTICLE

A new approach for incorporating ^{15}N isotopic data into linear inverse ecosystem models with Markov Chain Monte Carlo sampling

Michael R. Stukel^{1,2*}, Moira Décima³, Thomas B. Kelly¹

1 Department of Earth, Ocean, and Atmospheric Science, Florida State University, Tallahassee, FL, United States of America, **2** Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, FL, United States of America, **3** National Institute of Water and Atmospheric Research (NIWA), Wellington, New Zealand

* mstukel@fsu.edu



OPEN ACCESS

Citation: Stukel MR, Décima M, Kelly TB (2018) A new approach for incorporating ^{15}N isotopic data into linear inverse ecosystem models with Markov Chain Monte Carlo sampling. PLoS ONE 13(6): e0199123. <https://doi.org/10.1371/journal.pone.0199123>

Editor: Suzannah Rutherford, Fred Hutchinson Cancer Research Center, UNITED STATES

Received: January 30, 2018

Accepted: June 3, 2018

Published: June 18, 2018

Copyright: © 2018 Stukel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The model code used in this manuscript are available on GitHub: <https://github.com/stukel-lab/N15-LIM>.

Funding: Funding to support this work came from two sources: A National Science Foundation grant to the California Current Ecosystem Long-Term Ecological Research Program (OCE-1637632). https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=11696. A National Oceanic and Atmospheric Administration's RESTORE Program Grant to M. Stukel (Project Title: Effects of nitrogen

Abstract

Oceanographic field programs often use $\delta^{15}\text{N}$ biogeochemical measurements and *in situ* rate measurements to investigate nitrogen cycling and planktonic ecosystem structure. However, integrative modeling approaches capable of synthesizing these distinct measurement types are lacking. We develop a novel approach for incorporating $\delta^{15}\text{N}$ isotopic data into existing Markov Chain Monte Carlo (MCMC) random walk methods for solving linear inverse ecosystem models. We test the ability of this approach to recover food web indices (nitrate uptake, nitrogen fixation, zooplankton trophic level, and secondary production) derived from forward models simulating the planktonic ecosystems of the California Current and Amazon River Plume. We show that the MCMC with $\delta^{15}\text{N}$ approach typically does a better job of recovering ecosystem structure than the standard MCMC or L_2 minimum norm ($L_2\text{MN}$) approaches, and also outperforms an $L_2\text{MN}$ with $\delta^{15}\text{N}$ approach. Furthermore, we find that the MCMC with $\delta^{15}\text{N}$ approach is robust to the removal of input equations and hence is well suited to typical pelagic ecosystem studies for which the system is usually vastly under-constrained. Our approach is easily extendable for use with $\delta^{13}\text{C}$ isotopic measurements or variable carbon:nitrogen stoichiometry.

Introduction

Reconstructing ecosystem structure and trophic flows through planktonic ecosystems is crucial for understanding fisheries production, pelagic biogeochemistry, and the response of each of these to a changing climate. However, quantitative investigation of energy transfer between plankton functional groups is hampered by methodological limitations in separating ecological groups and in making rate measurements on specific trophic levels. As a result, oceanographers often rely on mass-balance ecosystem models such as EcoPath [1] or Linear Inverse Ecosystem Models (LIM [2, 3]) to reconstruct trophic structure from sparse measurements. However, the paucity and poor taxonomic resolution of common marine ecological measurements leaves such modeling approaches vastly under-constrained (e.g. [4]). Stable isotope

sources and plankton food-web dynamics on habitat quality for the larvae of Atlantic bluefin tuna in the Gulf of Mexico) under federal funding opportunity NOAA-NOS-NCCOS-2017-2004875. <https://restoreactscienceprogram.noaa.gov/funded-projects/bluefin-tuna-larvae>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

signatures ($\delta^{15}\text{N}$ and $\delta^{13}\text{C}$) offer additional information on diet and trophic position but are limited by difficulties in physically separating plankton functional groups with overlapping size [5–7]. Additionally, in contrast to terrestrial or estuarine systems where bulk stable isotopes of different primary producers are very different (e.g. depending on C3 or C4 plant metabolism), such differing bulk values have not been observed within plankton size classes [8], although arguably this could also be due to the size overlap of organisms with different isotope values [9]. Approaches capable of combining isotopic data and mass-balance approaches are thus clearly needed [10].

Although the frequency of *in situ* pelagic ^{15}N isotopic measurements is increasing, these data are seldom incorporated into inverse ecosystem models for two related reasons. First, although $\delta^{15}\text{N}$ measurements are tractable for several of the inputs and outputs to the planktonic ecosystem (e.g. metazoan zooplankton and larger organisms, sediment trap-based sinking material, NO_3^-), few measurements are made of intermediate compartments in the ecosystem, such as different phytoplankton or protozoan taxa, due to the methodological difficulty of these measurements [9], although particulate organic matter (POM) which includes a combination of the mentioned compartments in addition to detritus and bacteria, is often measured as well. This paucity of information on lower trophic levels would necessitate the use of a variable ^{15}N isotopic ratio for these taxa, which makes it impossible to linearize the equations. To date, studies incorporating variable isotopic ratios into LIM have included, at most, three compartments with unknown ^{15}N [11, 12]. Instead, $\delta^{15}\text{N}$ measurement results are typically used in stable isotope mixing models that often overly simplify the linkages in a planktonic ecosystem [13, 14]. Such models can, for instance, determine the relative role of nitrate uptake and nitrogen fixation in supporting pelagic new production if $\delta^{15}\text{N}\text{-NO}_3^-$ and $\delta^{15}\text{N}$ -exported material are measured [15–17]. However, this approach relies on the assumption that sediment trap caught material is representative of the isotopic signature of all material exported from the system, and thus neglects export mediated by larger organisms that are often enriched in ^{15}N . Similarly, attempts to estimate fish trophic levels often use simple linear mixed model approaches that assume that specific baseline consumers (e.g. crustacean zooplankton or filter-feeding benthic organisms) are strictly herbivorous (i.e. at a trophic level of two) [18–21] despite the facts that there can be multiple trophic steps within protists and the baseline consumers often feed efficiently on microzooplankton. Powerful new stable isotope models based on Bayesian statistical methods provide an additional approach for incorporating stable isotope measurements with prior knowledge of organism diet and isotopic fractionation [22–24]. However, these approaches are often developed for analyzing only a single trophic step and assume that the stable isotope signature of all dietary sources can be measured (but see [25] for one notable exception). Pacella et al. [26] combined Bayesian and LIM methods by using stable isotope analysis in R (SIAR) to incorporate dual isotopic signatures ($\delta^{13}\text{C}$ and $\delta^{15}\text{N}$) and constrain the dietary compositions of organisms from a seagrass ecosystem. These dietary compositions then provided powerful additional constraints for the LIM model. However, this approach can only be used for taxa for which the isotopic signatures of all dietary sources are known and hence is not well suited to pelagic ecosystems. An approach capable of linking a trophic and biogeochemical perspective (and assimilating measurements from both disciplines) would clearly be useful.

In this manuscript, we develop a framework for simultaneously incorporating $\delta^{15}\text{N}$ measurements and ecological rate measurements into a LIM while using Markov Chain Monte Carlo (MCMC) sampling to thoroughly sample the under-determined solution space. A scheme previously developed by van Oevelen et al. [12] for incorporating $\delta^{15}\text{N}$ data into LIMs relied on the L_2 minimum norm (L2MN) approach. However, inherent biases in this L2MN approach [27], suggest that methods reliant on it may be less accurate than approaches using

the more recently developed MCMC LIM methodology [28, 29]. Indeed, the L2MN approach has since been shown to be a less robust predictor of ecosystem flows than the MCMC approach [28, 30]. Studies incorporating nitrogen or carbon isotopic information from benthic ecosystems into the MCMC approach have either assumed that the isotopic signature of all compartments within an ecosystem are known [31] or used enriched stable isotope tracer additions to trace mass flow within the ecosystem [32, 33]. Since neither of these approaches is feasible for pelagic systems in which methodological considerations make it impossible to determine the isotopic signatures of many ecosystem compartments, we develop a new approach for incorporating isotopic information into the existing MCMC LIM framework. We follow the approach of Vézina and Pahlow [34] and use forward ecosystem models to generate simulated ecosystem flows; then use these simulated ecosystems to test the efficacy with which our new approach (MCMC+ ^{15}N), the Van Oevelen approach (L2MN+ ^{15}N), and the traditional MCMC and L2MN approaches recover key ecosystem parameters when supplied only with the types of data that are typically collected at sea.

Methods

We use two different forward models (NEMURO and DIAZO) and two separate sets of boundary conditions for each to create four different sets of fully constrained ecosystem flows. Because LIM models are designed to investigate the static, mass-balanced ecosystem structure, we run each model to steady-state in a simple one-box ecosystem configuration. We then determine input measurements—including ecosystem rate measurements (phytoplankton net primary production, nitrate uptake, mesozooplankton grazing, and sinking particle flux) and $\delta^{15}\text{N}$ values of allochthonous nitrate sources, DOM, mesozooplankton, and sinking particles—that are representative of the types of measurements that can be made in the field. We use these input measurements to constrain LIMs solved using four different approaches: standard L2MN and MCMC approaches and L2MN and MCMC approaches that incorporate $\delta^{15}\text{N}$ data. We then evaluate the ability of these LIM approaches to recover key ecosystem parameters related to nitrogen biogeochemistry (the relative proportion of phytoplankton production supported by new nitrate or nitrogen fixation) and plankton trophic dynamics (mesozooplankton mean trophic level and secondary production). We also conduct additional tests in which fewer ecosystem rate measurements are used as inputs to the inverse model. The following sections explain the forward models used and the four LIM approaches tested.

Forward models

To generate test datasets with known ecosystem flows (Table 1) we used two published nitrogen-based ecosystem models (NEMURO and DIAZO) with distinctly different model structure and behavior. NEMURO was developed to investigate zooplankton secondary production in the North Pacific for inclusion into ecosystem-based fisheries models [35, 36]. It contains three nutrient pools (nitrate, ammonium, and silicic acid) and two phytoplankton taxa (diatoms and small phytoplankton). It also has a relatively complex zooplankton community including protozoans, small herbivorous mesozooplankton, and large “predatory” mesozooplankton (that are actually omnivorous in the model). We configured the NEMURO model to run in a simple system designed to approximate the ecological dynamics of the euphotic layer in the southern California Current Ecosystem (CCE). Two versions of NEMURO were run, representing the coastal upwelling and offshore oligotrophic regions of the CCE. In the coastal configuration a 20-m deep euphotic zone was parameterized with upwelling rates of 1 m d^{-1} , while in the offshore region a 100-m euphotic zone was parameterized with 0.1 m d^{-1} upwelling. We parameterized NEMURO using the phytoplankton values determined by Li et al. [37]

Table 1. Comparison of the structures of the two forward (dynamical) models (NEMURO and DIAZO) and the inverse model (LIM).

	Nutrients	Phytoplankton	Zooplankton	Bacteria	Diazotrophy	Configurations
NEMURO	NO ₃ NH ₄ Si	Small ^{S,M} Large ^{M,P}	Small (Protists) Mesozooplankton Pred. Mesozoo.	Implicit	No	Coastal Offshore
DIAZO	Nitrogen Phosphorus Silica	Trichodesmium ^P Unicellular Diazotrophs ^P Cyanobacteria ^P Diatoms ^{M,P} DDA ^{M,P}	Protists Mesozooplankton	Implicit	Yes	Coastal Mesohaline
LIM	NO ₃ NH ₄	Small Phytoplankton ^{N,μ} Large Phytoplankton ^{μ,M}	Nanoflagellates (μ)Microzooplankton Mesozooplankton	Explicit	Yes	L2MN L2MN+ ¹⁵ N MCMC MCMC+ ¹⁵ N

Under phytoplankton categories, the superscripts denote which zooplankton groups consume the phytoplankton group. The 4 LIM configurations (L2MN, L2MN+¹⁵N, MCMC, MCMC+¹⁵N) are each run 4 times to simulate the 4 forward model configurations (NEMURO-Coastal, NEMURO-Offshore, DIAZO-Coastal, DIAZO-Offshore).

<https://doi.org/10.1371/journal.pone.0199123.t001>

from *in situ* rate measurement experiments conducted on CCE Long-Term Ecological Research cruises.

The DIAZO model was developed to investigate diazotroph growth and mortality terms in the Amazon River Plume [38]. As a result it has a diverse phytoplankton community including cyanobacteria, unicellular microbial diazotrophs, diatoms, diatom-diazotroph assemblages, and *Trichodesmium*. DIAZO has three nutrient pools (dissolved inorganic nitrogen, dissolved inorganic phosphorus, and silicic acid) and two zooplankton compartments (protozoans and mesozooplankton). Unlike NEMURO, DIAZO also allows protozoans to feed on themselves, thus simulating the longer protozoan food webs found in oligotrophic regions. We configured DIAZO to run in two 1-D systems: a low salinity coastal plume region (with high nutrient concentrations derived from the Amazon River outflow) and a mesohaline region with depleted NO₃⁻ but high Si and PO₄⁺ that is believed to be an ideal habitat for diazotrophs [39, 40].

In the DIAZO and NEMURO models, mesozooplankton mortality equations simulate predation by un-modeled higher trophic levels and other mortality terms and return the consumed mesozooplankton nitrogen to the detritus, dissolved organic nitrogen (DON), and dissolved nutrient pools. Since export mediated by sinking mesozooplankton carcasses or un-modeled higher trophic levels (including active transport and fecal pellets produced by fish and squid) is not typically captured in sediment traps, we modified both models such that mesozooplankton biomass consumed by this higher trophic level consumption was simply removed from the model (i.e. mesozooplankton secondary production is a sink term in the planktonic ecosystem).

Since our goal was to investigate the utility of δ¹⁵N isotopic data for inverse modeling reconstructions of pelagic food webs, we added a nitrogen isotopic model to NEMURO and DIAZO. The isotopic model was based on Yoshikawa et al. [41] and adds extra state variables representing the concentration of ¹⁵N in each nitrogen-containing model compartment (all living compartments, nutrients, DON, and detritus). Following Yoshikawa et al. [41], phytoplankton take up NO₃⁻ with a ¹⁵N fractionation factor (ε_{NO3}) of -5‰ and NH₄⁺ (and dissolved inorganic nitrogen (DIN) in the DIAZO model) with ε_{NH4} = -10‰. Zooplankton excretion and egestion are accompanied by fractionation factors of ε_{exc} = -5‰ and ε_{eg} = -2‰, respectively, while remineralization of detritus or DON to NH₄⁺ has ε_{rem} = -1‰. We assume that

nitrogen fixation (not included in [41]) introduces nitrogen to the ecosystem with a ^{15}N isotopic ratio equal to that of atmospheric nitrogen.

The DIAZO and NEMURO models with ^{15}N were both run to steady state using a 15-minute time step. Our goal with these simple 1-D models was not to faithfully simulate the full dynamics of the ecosystems modeled (although the steady state solutions agree reasonably with *in situ* measurements of phytoplankton size structure and growth rates and zooplankton grazing structure measured in the California Current Ecosystem [42–44] or Amazon River Plume [39, 45, 46]). Rather, we wanted to produce simple steady-state inputs representing four distinct ecosystem states (CCE coastal and offshore; Amazon Plume coastal and mesohaline) using models that differed substantially from each other and from the underlying ecosystem structure that we use for the inverse model. This replicates the difficulty encountered when attempting to model an *in situ* ecosystem for which the true structure of the system is unknown and hence the inverse model structure cannot be assumed to perfectly match the *in situ* plankton functional groups.

From these 4 forward simulations we determine steady-state rate values (net primary production, NO_3^- uptake, mesozooplankton grazing, and sediment trap-derived export), standing stocks (cyanobacteria biomass, diatom biomass, mesozooplankton biomass), and $\delta^{15}\text{N}$ values (of exogenous NO_3^- entering the ecosystem, mesozooplankton, DOM, and sinking detritus) that are representative of measurements that can be readily made in the field. We use these values as input measurements for the LIM approaches (see below). We then assess each LIM approaches' ability to recover 15 withheld forward model results for each forward model configuration: N_2 fixation, total phytoplankton nitrogen uptake, cyanobacteria net primary production (NPP), diatom NPP, protozoan grazing rate (on phytoplankton), mesozooplankton carnivory (on other mesozooplankton), protozoan carnivory (on other protists), mesozooplankton secondary production, protozoan gross growth efficiency (GGE), mesozooplankton GGE, protozoan trophic level (TL), mesozooplankton TL, herbivorous food chain (the percentage of total NPP that mesozooplankton consume through direct grazing of phytoplankton), multivorous food chain (proportion of NPP that is consumed in a food chain including multiple zooplankton groups before contributing to higher trophic levels), microbial loop food web (ratio of bacterial remineralization of DOM to total NPP). Note that bacterial remineralization is implicit in both forward models and explicit in the LIM models.

Linear Inverse Model (LIM) ecosystem structure

We use a simple LIM ecosystem structure that is characteristic of many LIM models of the planktonic ecosystem and borrows elements from Jackson & Eldridge [47] and Richardson et al. [48]. Our LIM includes 5 living compartments: small phytoplankton (SPHY), large phytoplankton (LPHY), heterotrophic nanoflagellates (HNF), microzooplankton (MIC), and mesozooplankton (MES). It also includes 4 non-living compartments: ammonium (NH_4), nitrate (NO_3), detritus (DET), and dissolved organic matter (DOM). All model flows are measured in nitrogen currency ($\text{mmol N m}^{-2} \text{d}^{-1}$). Phytoplankton production is supported by uptake of NO_3 and NH_4 , and nitrogen fixation. Heterotrophs graze on phytoplankton (MES are assumed to be incapable of grazing on SPHY) as well as on other heterotrophs smaller than themselves. Mesozooplankton are also assumed to be capable of carnivory on themselves (though other grazers do not feed on themselves). All living taxa lose nitrogen to DOM (excretion) and to DET (defecation by grazers, mortality by phytoplankton). DET is remineralized to DOM and DOM is remineralized to NH_4 . Inputs of nitrogen to the ecosystem (upwelled or advected NO_3 and nitrogen fixation) are balanced by sinking DET and loss of MES to higher trophic levels (HTL). There are a total of 35 ecosystem flows (Figs 1 and 2).

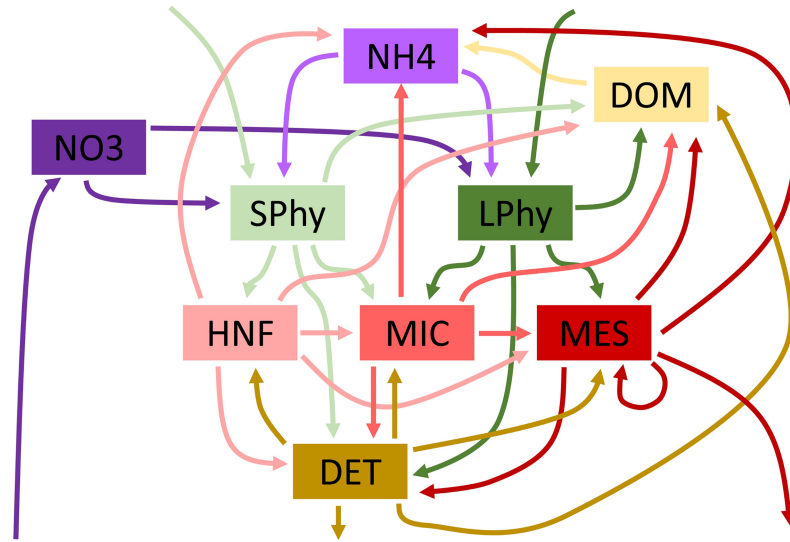


Fig 1. Schematic diagram showing LIM ecosystem flows. Compartments are nitrate (NO₃), ammonium (NH₄), small phytoplankton (SPhy), large phytoplankton (LPhy), heterotrophic nanoflagellates (HNF), microzooplankton (MIC), mesozooplankton (MES), detritus (DET), and dissolved organic matter (DOM).

<https://doi.org/10.1371/journal.pone.0199123.g001>

In the LIM we solve a series of systems of equations including:

$$Ex = f \tag{1}$$

which quantifies mass balance constraints on each model compartment (9 equations).

$$Ax \approx b \tag{2}$$

which quantifies measurement constraints with associated uncertainty (primary production, nitrate uptake, mesozooplankton grazing, and sediment trap flux). We assume that measurement uncertainty in these parameters is $\pm 10\%$. For the MCMC+¹⁵N and L2MN+¹⁵N models we also include ¹⁵N mass balance constraints. These constraints are included in the approximate equations because, while mass balance must hold, we assume nitrogen fractionation factors included in these equations are not exactly known. We assume that uncertainty in these mass balance equations is equal to 10% of the sum of the mass flow through the compartment multiplied by the expected isotopic fractionation between compartments.

$$Gx \geq h \tag{3}$$

which quantifies *a priori* assumed greater than / less than constraints on organisms within the ecosystem (e.g. gross growth efficiency of grazers varies between 10 and 40%). A full list of greater than / less than equations can be found in the [S1 Supplementary Text](#).

L2MN and MCMC model approaches

With 35 unknown flows and only 9 exact equalities and 4 approximate equalities (and an additional 10 approximate equalities for the LIM solutions including ¹⁵N), the model is substantially under-constrained. Hence there are an infinite number of solutions that can satisfy Eqs 1, 2, and 3 (unless some of the equations and inequalities are inconsistent). To choose amongst this infinite number of solutions to the inverse problem, the study that originally brought LIM methodology to aquatic ecology used an L2MN approach [3, 49]. This approach requires that the solution must satisfy Eq 1 and In Eq 3. The L2MN approach then chooses the solution that

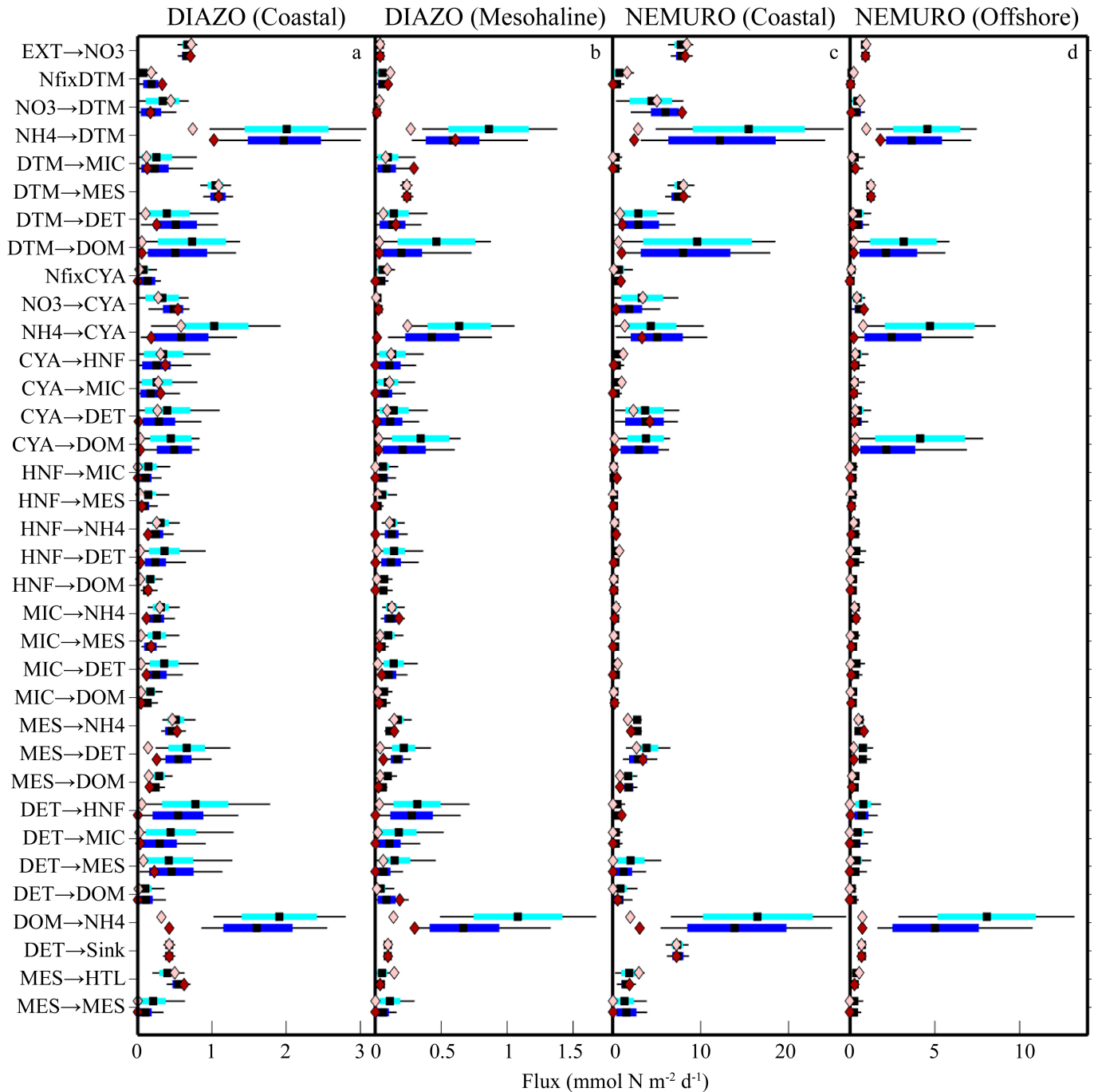


Fig 2. LIM solutions for DIAZO (Coastal) model (a), DIAZO (Mesohaline) model (b), NEMURO (Coastal) model (c), and NEMURO (Offshore) model (d). Units are mmol N m⁻² d⁻¹. Light and dark box plots show 95% confidence intervals, quartiles, and mean for MCMC and MCMC+¹⁵N, respectively. Light and dark diamonds show values determined by L2MN and L2MN+¹⁵N, respectively.

<https://doi.org/10.1371/journal.pone.0199123.g002>

minimizes Eq 2 and if there are multiple solutions that have zero residual error in Eq 2, the L2MN approach chooses the solution that minimizes the sum of squared flows through the ecosystem. To solve the inverse problem we used the function lsei contained in the R Package limSolve [50].

While the L2MN approach has been widely used, there is no *a priori* reason to assume that the solution that minimizes nitrogen (or carbon) flow through the ecosystem is the most appropriate solution. As an alternate approach, Kones et al. [28, 29] developed a Markov Chain Monte Carlo (MCMC) random walk method that uses a Metropolis algorithm to sample throughout the viable solution space. The MCMC approach initiates with a solution (x_0) that is known to solve Eqs 1 and 3. A new proposed sample (x_1) is then drawn from a random jump through the region constrained by Eqs 1 and 3. The residual error of x_1 with respect to Eq 2 is then compared to the residual error of x_0 to determine whether the new solution should be accepted. If x_1 is accepted the process is iterated to determine another solution. If not, the process is repeated from x_0 . In this way a constrained random walk is performed through the solution space. This iteration procedure produces a target distribution of solutions satisfying Eqs 1–3. Summary statistics (mean, standard deviation, confidence intervals) for each ecosystem flow can then be calculated. For more details, we refer readers to Van den Meersche et al. [51] and van Oevelen et al. [2]. The MCMC mean solution has been shown to more accurately estimate in situ ecosystem measurements that are withheld as inputs to the model [30, 52]. It also avoids the undesirable tendency that the L2MN approach has for choosing solutions that are extreme values of the possible solution space (i.e. the L2MN approach often selects solution sets that rest on one of the hyperplanes formed by $Gx \geq h$). We implemented the MCMC approach using the R function `xsample` [51].

L2MN+¹⁵N model approach

Incorporating ¹⁵N into a LIM is essentially equivalent to incorporating any other form of second currency into the model (e.g. using carbon and nitrogen to constrain the model). Such models are a trivial extension of previous LIM approaches if the stoichiometry (C:N or $\delta^{15}\text{N}$) of all model compartments is known or assumed. However, in most pelagic ecosystem studies the $\delta^{15}\text{N}$ (or C:N) values of many components of the ecosystem cannot be measured. In a study focused on benthic ecosystems, van Oevelen et al. [12] developed an approach for incorporating $\delta^{13}\text{C}$ measurements into a LIM model for which the $\delta^{13}\text{C}$ of three components of the ecosystem were unknown. Since trophic fractionation of $\delta^{13}\text{C}$ is minimal, they incorporated linear mixing equations into the equality constraints with the following form:

$$\delta^{13}C_j = \frac{\sum_i (\delta^{13}C_i \times flow_{i \rightarrow j})}{\sum_i flow_{i \rightarrow j}} \tag{4}$$

They then conducted a grid search ($\pm 0.1 \delta^{13}\text{C}$) calculating the L2MN solution for every reasonable combination of $\delta^{13}\text{C}$ values for the 3 compartments with unknown $\delta^{13}\text{C}$. The chosen solution was the solution that minimized the residual error and (if multiple solutions had zero residual error) the sum of squared ecosystem flows.

It was necessary to adapt this approach to work with $\delta^{15}\text{N}$, because trophic fractionation significantly affects the $\delta^{15}\text{N}$ values of a compartment, as they are determined not only by their sources of nitrogen but also by the degree of fractionation of the loss terms from that compartment. Thus we constructed a system of $\delta^{15}\text{N}$ mass balance equations for the flow of ¹⁵N into and out of each compartment, e.g. for NO₃:

$$R_{upno3} \times \text{EXT} \rightarrow \text{NO}_3 - (R_{no3} \times \alpha_{no3}) \times \text{NO}_3 \rightarrow \text{CYA} - (R_{no3} \times \alpha_{no3}) \times \text{NO}_3 \rightarrow \text{DTM} \\ = 0 \tag{5}$$

where R_{upno3} and R_{NO3} are the ratio of ¹⁵N to total N for allochthonous nitrate entering the ecosystem and euphotic zone nitrate, respectively, α_{no3} is the isotopic fractionation coefficient

for nitrate uptake and is equal to $\exp(\epsilon_{\text{no}_3})$, $\text{EXT} \rightarrow \text{NO}_3$ is the ecosystem flow corresponding to allochthonous nitrate entering the ecosystem, and $\text{NO}_3 \rightarrow \text{CYA}$ and $\text{NO}_3 \rightarrow \text{DTM}$ are the ecosystem flows corresponding to nitrate uptake by CYA and DTM. The full set of ^{15}N mass balance equations can be found in Appendix 1. These equations were incorporated into the approximate equations (Eq 2), because while mass balance must hold, we assume that fractionation terms are uncertain.

We included the $\delta^{15}\text{N}$ values of allochthonous nitrate entering the ecosystem, euphotic zone DOM, zooplankton, and sinking detritus as measured inputs given to the LIM, because these measurements can be readily made in the field (e.g. [53–56]). This left 6 model compartments (NO_3 , NH_4 , CYA , DTM , HNF , and MIC) for which $\delta^{15}\text{N}$ was unknown. We conducted a grid-search through this 6-dimensional grid space testing all realistic parameter ranges (with 0.25 ‰ step size), solving the L2MN for each $\delta^{15}\text{N}$ parameter set. The solution set with the lowest residual norm ($\sigma^{-2}(Ax-b)^T(Ax-b)$) was selected as the L2MN+ ^{15}N solution. We note, however, that a slightly better solution might be found with greater discretization of the tested $\delta^{15}\text{N}$ values although computational power limited the step size we could use (decreasing the step size from 0.25 to 0.1 would have required computing the L2MN ~20 billion times rather than ~100 million times). Use of the L2MN+ ^{15}N approach with models containing additional compartments will likely require the use of a gradient-based variational approach rather than a full grid search.

MCMC+ ^{15}N model approach

The MCMC approach has many desirable qualities, not least of which is the fact that it allows computation of model uncertainty resulting from both uncertainty in the inputs to the model and the inherent under-determinacy of the system [28]. The MCMC is also a more efficient sampler of the parameter space than a full grid search and we took advantage of this by including varying $\delta^{15}\text{N}$ values in the MCMC search algorithm. These unknown $\delta^{15}\text{N}$ values are stored in a new vector (∂_U), while the known (measured) $\delta^{15}\text{N}$ values are stored in a vector (∂_K). The $\delta^{15}\text{N}$ values are incorporated into the approximate equations (Eq 2) using the same mass balance constraints as for the L2MN+ ^{15}N (Appendix 1). The matrix A is thus now a function of ∂_U and ∂_K . During the MCMC algorithm when a new proposed sample (x_1) is determined with a random jump from the previous solution vector (x_0), we also propose a new set of $\delta^{15}\text{N}$ values ($\partial_{U,1}$) by performing a random jump from the prior set of $\delta^{15}\text{N}$ values ($\partial_{U,0}$). We then call a function that recalculates the matrix A_1 as a function of $\partial_{U,1}$ and ∂_K . We then proceed (as for the standard MCMC approach) to decide whether to accept or reject x_1 and $\partial_{U,1}$ based on the ratio of $p(x_1, A_1)/p(x_0, A_0)$. If the values are accepted, both x_1 and $\partial_{U,1}$ are appended to the overall solution. If not, they are rejected and the process is repeated from x_0 and $\partial_{U,0}$. This approach thus generates a series of solutions that satisfy the equality constraints (Eq 1) and inequality constraints (Eq 3), while approximately satisfying the input measurements and $\delta^{15}\text{N}$ mass balance equalities codified in Eq 2 and Appendix 1. Model code was written in the open source language R (3.3.2) to take advantage of existing algorithms in the `limSolve` package [57] and can be downloaded from GitHub at: <https://github.com/stukel-lab/N15-LIM>.

Model runs with less constraint

Due to the difficulty of measuring pelagic ecosystem rates in situ, most LIM studies have a lower ratio of measurement constraints to unknown flows than our base LIM studies. As a result, we conducted tests in which we withheld either single measurements or pairs of measurements. These simulations were conducted by sequentially withholding a single

measurement constraint (nitrate uptake, mesozooplankton grazing, or sediment trap-derived export) or pairs of those measurements. Primary productivity was never withheld as a measurement constraint, because it has been measured in every pelagic LIM study that we have encountered.

Statistical analysis

To compute 95% confidence intervals for derived variables (e.g. trophic level) from MCMC and MCMC+ ^{15}N approaches, we computed the derived variable for each individual solution and took the 95% confidence intervals of this distribution of derived variables. We depict these values as whisker plots showing means, quartiles, and 95% confidence intervals. For an objective comparison of the efficacy with which each LIM approach depicted the underlying “true” ecosystem structure we compared LIM estimates for each of the withheld results in [Table 2](#). These varied ecosystem indices could be similarly calculated for both forward models and the LIM ecosystem reconstructions and together gave a composite snapshot of ecosystem structure. To ensure that each of the 15 indices was given equivalent weight, we first pooled results for each individual index (e.g. all estimates of N_2 fixation from the four forward model runs and from the 112 LIM runs (4 forward models \times 4 approaches \times 7 input configurations)). We then used a two-parameter Box Cox transformation (R function `boxcofit`) to normalize this pooled data, subtracted off the mean, and divided by the standard deviation. Thus values for all indices were approximately normal, with a mean of 0 and standard deviation of 1. To determine a composite index that assessed the overall effectiveness with which any LIM model run recovered the underlying “true” values from the forward model, we computed the sum of squared errors (SSE) by subtracted the “true” value from the LIM prediction and summing the square of this value for all 15 indices. To visualize this data in two dimensions we used non-metric multi-dimensional scaling using the `mdscale` function in Matlab. Dissimilarities used as inputs for `mdscale` were calculated from Matlab function `pdist`.

Results and discussion

Forward model results

We used two models (DIAZO and NEMURO), each run twice to simulate different ecosystem conditions, to develop four ecosystem snapshots for use as inputs to the inverse modeling algorithms ([Table 2](#)). The DIAZO model was configured to simulate conditions in the coastal and mesohaline region of the Amazon River Plume. In the coastal region, nutrients derived from the Amazon River ($0.8 \mu\text{mol L}^{-1}$ phosphate, $32 \mu\text{mol L}^{-1}$ silicic acid, and $8.5 \mu\text{mol L}^{-1}$ nitrate with a riverine $\delta^{15}\text{N}$ value of 7.7) support a moderately productive ecosystem. Diatoms fueled 70% of net primary production and this diatom-dominance was reflected in the zooplankton community structure with mesozooplankton responsible for 60% of total grazing and maintaining a biomass 4 times greater than protistan zooplankton. These mesozooplankton were at a trophic level of 2.15 (reflecting a predominantly diatom diet) and had a $\delta^{15}\text{N}$ value of 6.10.

In the mesohaline region, where diatom-diazotroph assemblages are expected to be common, the DIAZO model predicted substantial N_2 fixation, with N_2 fixation supporting 74% of the total new production in the system. Net primary production was reduced relative to the coastal region ($0.70 \text{ mmol N m}^{-2} \text{ d}^{-1}$, compared to $2.17 \text{ mmol N m}^{-2} \text{ d}^{-1}$). Cyanobacteria:diatom biomass was more closely balanced in the mesohaline region and a substantial portion of the diatom community was comprised of diatom-diazotroph assemblages. Mesozooplankton had a similar trophic level to that in the coastal region (2.21) although they comprised a lower proportion of total grazing (40%) and had a substantially lower $\delta^{15}\text{N}$ signature of 5.08 reflecting the increased importance of N_2 fixation in this region.

Table 2. Model results from the DIAZO and NEMURO models (at steady state).

Model results used as inputs to inverse model	DIAZO (Coastal)	DIAZO (Mesohaline)	NEMURO (Coastal)	NEMURO (Oligotrophic)
Net Primary Production (mmol N m ⁻² d ⁻¹)	2.17	0.70	13.61	2.42
NO ₃ ⁻ Uptake (New Production) (mmol N m ⁻² d ⁻¹)**	0.72*	0.04*	8.44	0.95
Mesozooplankton Grazing (mmol N m ⁻² d ⁻¹)	1.09	0.24	8.11	1.25
Sediment Trap Export (mmol N m ⁻² d ⁻¹)	0.42	0.10	7.28	0.69
Cyanobacteria Biomass (mmol N m ⁻²)	1.55	1.21	12.12	14.65
Diatom Biomass (mmol N m ⁻²)	2.57	1.63	34.58	10.94
Mesozooplankton Biomass (mmol N m ⁻²)	6.38	2.14	18.38	17.99
Temperature (°C)	28.00	28.00	12.00	14.00
δ ¹⁵ N exogenous NO ₃ ⁻	7.70	13.25	5.70	5.70
δ ¹⁵ N Mesozooplankton	6.10	5.08	7.68	7.23
δ ¹⁵ N DOM	4.20	2.95	4.87	4.66
δ ¹⁵ N Sinking Detritus	3.98	2.88	4.42	4.82
Model results withheld from inverse model				
New Production (N ₂ Fixation) (μmol N m ⁻² d ⁻¹)	3.11	100.48	0	0
Total N Uptake (mmol N m ⁻² d ⁻¹)	3.09	0.96	17.16	3.53
Cyanobacteria Net Primary Production (mmol N m ⁻² d ⁻¹)	0.48	0.25	2.51	1.33
Diatom Net Primary Production (mmol N m ⁻² d ⁻¹)	1.69	0.45	11.10	1.09
Protozoan Grazing Rate (mmol N m ⁻² d ⁻¹)	0.73	0.36	0.53	0.74
Mesozooplankton Carnivory (MES→MES) (mmol N m ⁻² d ⁻¹)	0	0	1.89	0.24
Protozoan Carnivory (HNF→MIC) (mmol N m ⁻² d ⁻¹)	0.11	0.07	0	0
Mesozooplankton Secondary Production (mmol N m ⁻² d ⁻¹)	0.30	0.04	1.16	0.26
Protozoan GGE	0.3	0.3	0.3	0.3
Mesozooplankton GGE	0.3	0.3	0.3	0.3
Protozoan Trophic Level	2.12	2.22	2.00	2.00
Mesozooplankton Trophic Level	2.15	2.21	2.46	2.37
Herbivorous Food Chain ¹ (Percent Total NPP)	0.50	0.34	0.23	0.27
Multivorous Food Chain ² (Percent Total NPP)	0.34	0.51	0.01	0.16
Microbial Loop Food Web ³ (Percent Total NPP)	0.80	0.83	0.19	0.28
δ ¹⁵ N Euphotic Zone NO ₃ ⁻	13.25**	11.96**	9.92	8.84
δ ¹⁵ N NH ₄ ⁺	13.25**	11.96**	12.80	13.60
δ ¹⁵ N Cyanobacteria	3.03	1.73	3.91	3.65
δ ¹⁵ N Diatoms	3.17	1.59	3.87	3.64
δ ¹⁵ N Protozoans	6.17	5.04	6.52	6.26

Upper portion of table shows results that were given as input to the inverse models. Lower portion shows results that were withheld and can be used as comparisons for the inverse models.

¹Herbivorous Food Chain is proportion of NPP routed through the direct phytoplankton → mesozooplankton → higher trophic levels food chain.

²Multivorous food chain is proportion of NPP that enters food chains that pass through multiple zooplankton before reaching higher trophic level.

³Microbial loop is proportion of NPP that is (implicitly) processed by bacteria through degradation of DOM to NH₄⁺.

*The DIAZO model has only one dissolved inorganic nitrogen compartment (DIN, rather than NO₃⁻ and NH₄⁺), hence new production that would be equivalent to nitrate uptake is calculated from the proportion of exogenous (non-recycled) DIN used by phytoplankton.

**The δ¹⁵N value of the single DIN pool in DIAZO is given for both NO₃⁻ and NH₄⁺.

<https://doi.org/10.1371/journal.pone.0199123.t002>

The NEMURO model was used to simulate plankton communities in the coastal upwelling region of the California Current Ecosystem and the oligotrophic North Pacific Subtropical Gyre. Both systems were supported by upwelled nutrients at a concentration of 10 μmol NO₃⁻

L^{-1} and $10 \mu\text{mol Si L}^{-1}$, but the upwelling rates differed substantially (1 m d^{-1} into a 30-m euphotic layer in the coastal region; 0.1 m d^{-1} into a 100-m euphotic layer in the oligotrophic region). The coastal region exhibited the highest net primary productivity ($13.6 \text{ mmol N m}^{-2} \text{ d}^{-1}$) and f -ratio (new production / total production = 49%) of our model runs. Diatoms were the dominant phytoplankters (65% of biomass; 82% of production) and mesozooplankton were responsible for nearly all grazing (94%). Despite their substantial grazing rates, mesozooplankton had a high trophic level of 2.46, reflecting the prevalence of mesozooplankton carnivory (unlike in the DIAZO model, the NEMURO model contains a predatory mesozooplankton class that can feed on other mesozooplankton).

The offshore oligotrophic NEMURO run had the lowest net primary productivity by volume of any of our model runs (although vertically integrated primary productivity was higher than in the DIAZO model runs, as we assumed a 10-m mixed layer for the Amazon Plume and a 100-m euphotic zone for the oligotrophic subtropical gyre). This low productivity was matched by cyanobacterial dominance of the phytoplankton community (57% of biomass; 55% of production). However, this cyanobacteria production did not translate into an important role for protozoans, which were responsible for 37% of the grazing. This high cyanobacteria / low protozoan condition is determined by the structure of NEMURO, which allows mesozooplankton to graze on both phytoplankton classes, but restricts protozoans to consuming only cyanobacteria. Mesozooplankton had a lower trophic level and $\delta^{15}\text{N}$ than in the coastal upwelling region, reflecting reduced rates of carnivory.

Inverse model ecosystem reconstructions

Results from the forward model run (primary production, nitrate uptake, mesozooplankton grazing, cyanobacteria biomass, diatom biomass, mesozooplankton biomass, and the $\delta^{15}\text{N}$ signatures of exogenous NO_3^- , mesozooplankton, DOM, and sinking detritus) were used to force LIM simulations using the L2MN, L2MN+ ^{15}N , MCMC, and MCMC+ ^{15}N approaches. To assess the accuracy of the nutrient dynamics in each of the LIM simulations, we compared the ratio of NO_3^- uptake to total nitrogen uptake and N_2 fixation to total nitrogen uptake (Fig 3, S1 Fig). A distinct difference was clear between the L₂ minimum norm approaches and the Monte Carlo approaches, with both MCMC approaches consistently underestimating the ratio of percent nitrogen taken up as new nitrate and the L2MN approaches typically overestimating nitrate uptake percentage. However, the addition of ^{15}N information consistently improved the MCMC+ ^{15}N approach relative to the standard MCMC approach. The mean percent error decreased from a 43% underestimate with the MCMC approach to a 27% underestimate with the MCMC+ ^{15}N approach. When comparing the LIM approaches' recovery of N_2 fixation rates all four approaches overestimated N_2 fixation for the DIAZO coastal run and both NEMURO model runs (which was unsurprising since the NEMURO model does not allow N_2 fixation). However, the L2MN approach (without ^{15}N) typically performed worse than the other three approaches, estimating that N_2 fixation was always between 8 and 13% of total nitrogen uptake (when the "true" value was negligible), while the other three approaches were comparable and suggested fractions of 1% - 9% (except for the L2MN+ ^{15}N approach with the DIAZO coastal model run, which predicted 15%). For the DIAZO mesohaline run, the "true" value from the forward model was 10% and the MCMC approaches slightly underestimated this value (7% for MCMC; 9% for MCMC+ ^{15}N), while the L2MN+ ^{15}N approach slightly overestimated it at 13% and the L2MN approach substantially overestimated it at 27%.

To determine the ability of the LIM approaches to reconstruct grazer dynamics, we compared the forward model values to trophic level and secondary production estimates from the LIM models (Fig 4, S2 Fig). Both MCMC approaches did a reasonable job of recovering

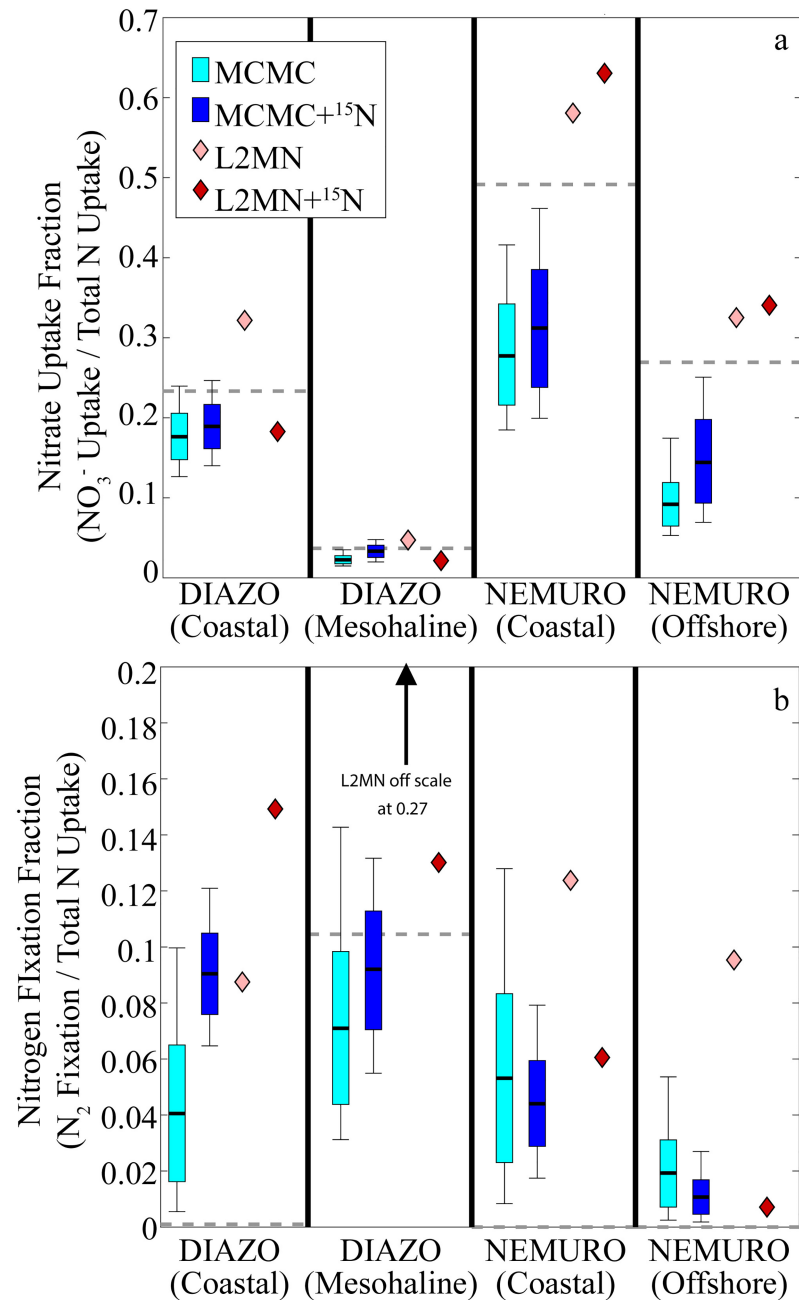


Fig 3. LIM NO₃⁻ Uptake (a) and N₂ Fixation (b). Both plots show fraction of total phytoplankton nitrogen supplied by respective process. Light and dark box plots show 95% confidence intervals, quartiles, and mean for MCMC and MCMC+¹⁵N, respectively. Light and dark diamonds show values determined by L2MN and L2MN+¹⁵N, respectively. Dashed gray line shows “true” value from the forward model run.

<https://doi.org/10.1371/journal.pone.0199123.g003>

mesozooplankton trophic levels for all model simulations (95% confidence intervals consistently bracketed the “true” values), although they were often high or low by ~0.2 trophic levels. However, the L2MN approaches were biased low, particularly with the NEMURO model, for which they predicted trophic levels ranging from 2.01–2.12, while the “true” values were 2.46 and 2.37. For mesozooplankton secondary production (which we define herein as the amount of mesozooplankton production that was consumed by higher trophic levels), we found that

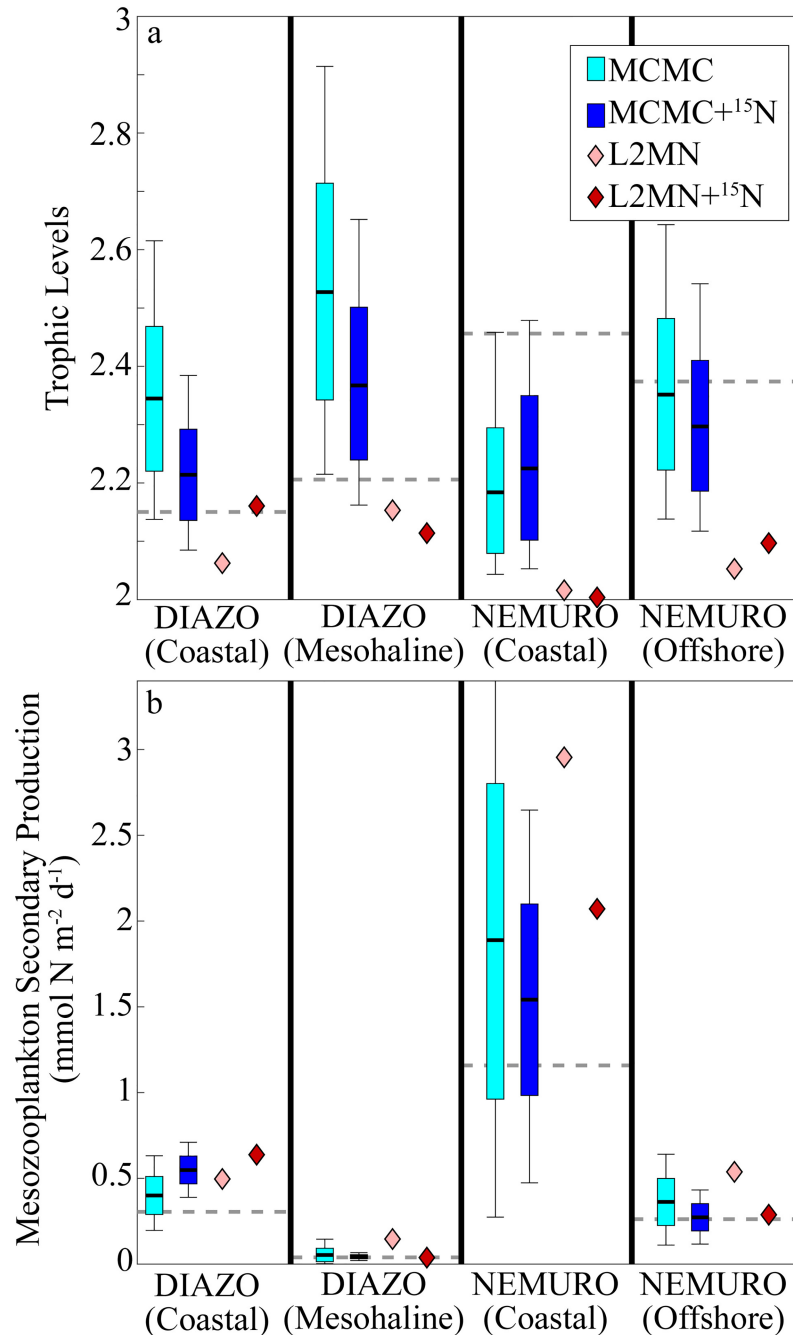


Fig 4. LIM Mesozooplankton trophic level (a) and secondary production (b). LIM secondary production is the flow from MES to unmodeled higher trophic levels. Light and dark box plots show 95% confidence intervals, quartiles, and mean for MCMC and MCMC+¹⁵N, respectively. Light and dark diamonds show values determined by L2MN and L2MN+¹⁵N, respectively. Dashed gray line shows “true” value from the forward model run.

<https://doi.org/10.1371/journal.pone.0199123.g004>

the MCMC, MCMC+¹⁵N, and L2MN+¹⁵N approaches did a good job of recovering the “true” values for the oligotrophic ecosystem states, but were biased slightly high for the coastal ecosystem states. By contrast, the standard L2MN approach consistently overestimated secondary production, at times predicting values that were greater than three times the expected value.

Inverse model performance with less constraint

The LIM structure used here is relatively simple and well-constrained compared to many inverse models that include greater taxonomic complexity amongst the plankton, higher trophic level compartments, or multi-layer ecosystems (e.g. [47, 48, 58]). As a result, the rank parameter for our LIM structure (which denotes the number of linearly independent equations in Eqs 1 and 3, and is equal to the number of unknown flows minus the number of constraining equations) is much lower than for many published ecosystem models. For comparison, the rank of the base L2MN structure of our model (35 ecosystem unknowns, 9 mass balance constraints, 4 measurement constraints) is 22, while the Antarctic ecosystem LIM of Sailley et al. [58] contained 48 unknowns, 10 mass balance constraints and only 2 measurement constraints (rank = 36) and the Equatorial Pacific model with size-fractionated detritus of Stukel and Landry [4] contained 62 unknowns, 12 mass balance constraints, and 8 measurement constraints (rank = 42). To determine the efficacy of the 4 LIM approaches when the system has less constraint, we recomputed the solutions while sequentially withholding one or two of our measurement constraints (with the exception of NPP, which has been used as an input for every pelagic ecosystem LIM that we have seen), thus yielding 6 additional LIM solutions for each approach (3 with a single measurement withheld, 3 with pairs of measurements withheld).

When assessing mesozooplankton dynamics (trophic level and secondary production), the MCMC+ ^{15}N model showed relatively little degradation in accuracy as the model became more underconstrained (Fig 5A–5D). In fact, for the two coastal simulations, trophic level estimates did not change substantially when measurements were withheld, but secondary production estimates decreased to more closely match the “true” value determined by the forward model run. For the mesohaline and offshore runs, the model performed slightly more poorly on average when measurements were withheld. Similar results were found when comparing nitrogen dynamics estimated with measurements withheld (Fig 5E–5H); results obtained by the MCMC+ ^{15}N approach were relatively insensitive to which measurements were used as inputs to the model. By contrast, the offset between the MCMC estimate and the “true” value tended to be a bit larger when more measurements were withheld, while solution sets from both L2MN approaches exhibited strong sensitivity to which measurements were used as inputs (error often switched from over- to under-estimates, or vice versa, depending on which measurements were withheld).

For a holistic comparison of the efficacy with which each LIM approach depicted the underlying “true” ecosystem structure we compared LIM estimates for each of the withheld results in Table 2. We computed the sum of squared errors (SSE) for each LIM model run with respect to the “true” values for each of these indices. This composite value reflects how well the LIM ecosystem structure captures the overall value. With respect to SSE, the MCMC+ ^{15}N approach performed better than the standard MCMC approach for 27 of the 28 model-measurement input pairings and was the best of the four approaches for 23 of the 28 scenarios. Compared to the other approaches, the MCMC+ ^{15}N approach also showed less deterioration in accuracy when measurements were withheld. With all input measurements included, the median SSE for the MCMC+ ^{15}N approach was 20.7 while it was 20.8 with one measurement withheld and 20.2 with two measurements withheld. The relative insensitivity of the MCMC+ ^{15}N approach to removal of measurements is also clear when the results are plotted using non-metric multi-dimensional scaling (NMDS, Fig 6). While the MCMC approaches tended to cluster together, the results from the L2MN approaches were highly sensitive to which measurements were used as inputs and often produced results on opposite sides of the NMDS plot. These highly divergent results arise from the L2MN approaches’ tendency to select solution sets that lie on the

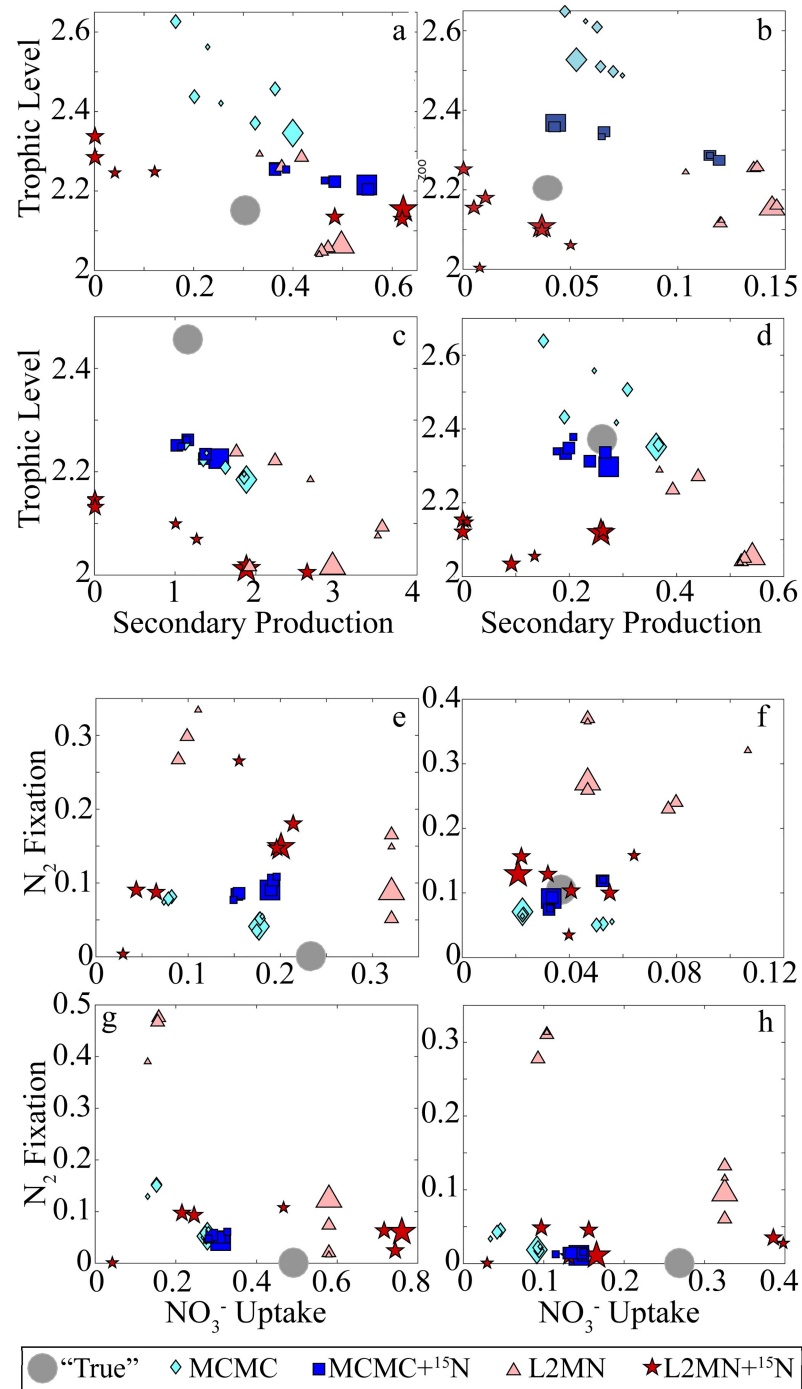


Fig 5. LIM accuracy when input measurements are withheld. Panels a-d show mesozooplankton secondary production ($\text{mmol N m}^{-2} \text{d}^{-1}$, x-axis) against mesozooplankton trophic level (y-axis). Panels e-g show NO_3^- uptake (fraction of total N uptake, x-axis) against N_2 fixation (fraction of total N uptake, y-axis). Gray circles indicate the "true" values from the forward models. Other symbols are MCMC (square), MCMC+ ^{15}N (diamond), L2MN (triangle), and L2MN+ ^{15}N (star). Symbol size reflects the number of measurements withheld as inputs for the inverse model (large is no measurements withheld, medium is one measurement withheld, small is two measurements withheld). a,e) DIAZO Coastal; b,f) DIAZO Mesohaline; c,g) NEMURO Coastal; d,h) NEMURO Offshore.

<https://doi.org/10.1371/journal.pone.0199123.g005>

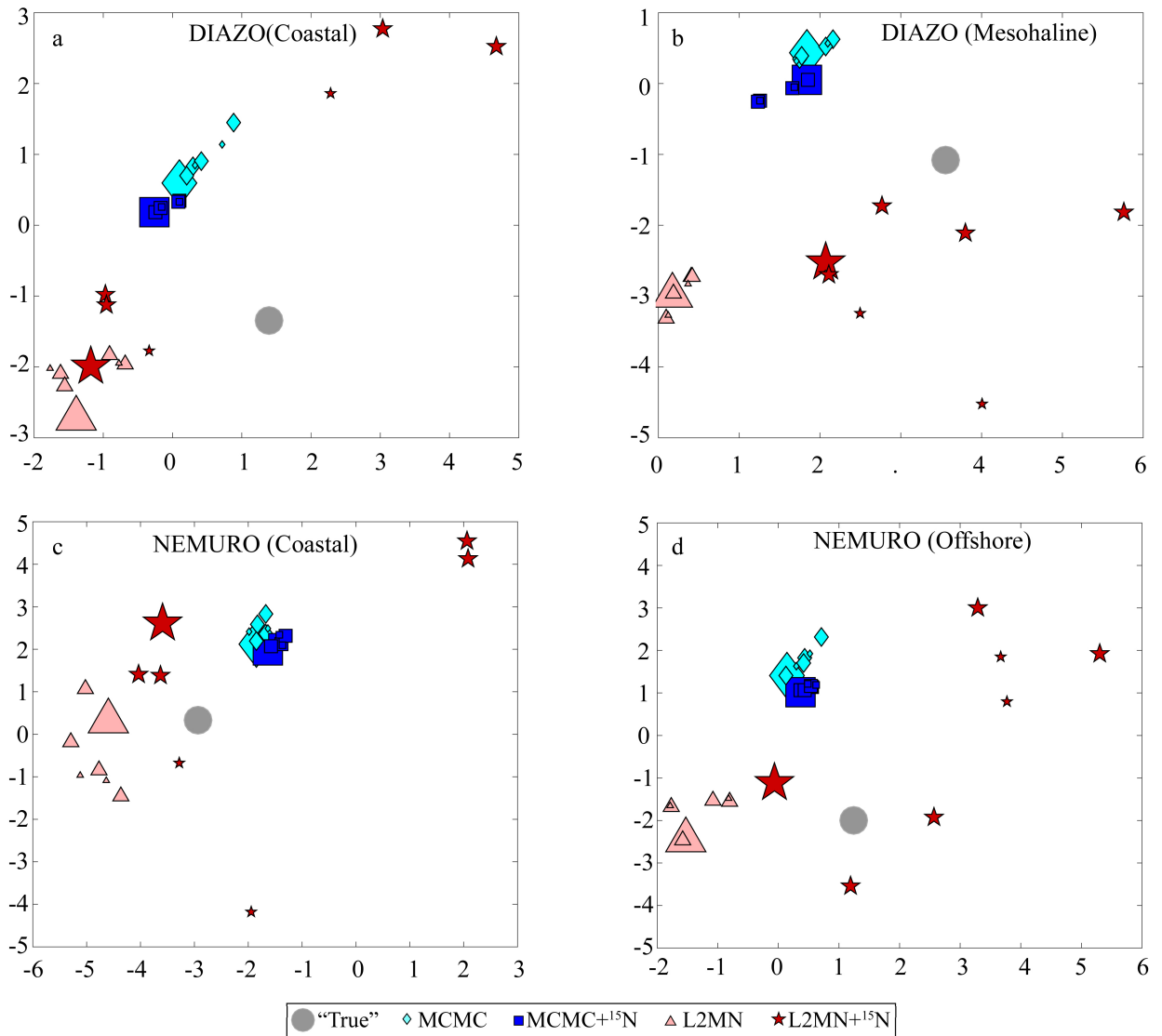


Fig 6. Non-metric multi-dimensional scaling (NMDS) plot showing model variability with respect to the 15 model results withheld from the inverse model in Table 2 (excluding $\delta^{15}\text{N}$ values). The minimized stress of the NMDS analysis was 0.16.

<https://doi.org/10.1371/journal.pone.0199123.g006>

edge of the possible solution space (e.g. gross growth efficiency, which can vary from 0.1 to 0.4 is usually placed at exactly 0.1 or 0.4).

$\delta^{15}\text{N}$ values

The MCMC+¹⁵N approach also did an excellent job of recovering the $\delta^{15}\text{N}$ signatures of model compartments (Fig 7). Most of the time, the model 95% confidence intervals bracketed the actual values and a Type 1 linear regression of the MCMC+¹⁵N predicted values regressed against the "true" values was statistically significant ($p < 0.001$) with a slope of 1.05, an intercept of 0.39, and an r^2 value of 0.90. The only major discrepancy between the MCMC+¹⁵N values and the "true" values arose with the $\delta^{15}\text{N}$ value of NO_3^- for the DIAZO model. This misfit was not unexpected as the DIAZO model has only one dissolved inorganic nitrogen state variable (the $\delta^{15}\text{N}$ value of which was used for the "true" value for both NO_3^- and NH_4^+). In most

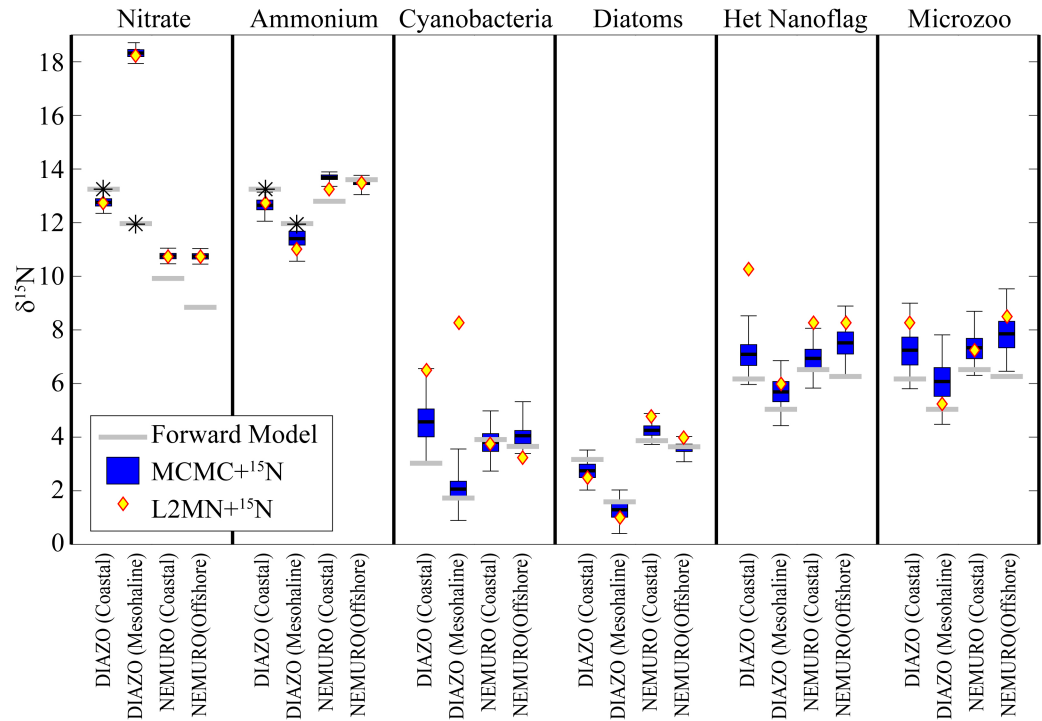


Fig 7. LIM model estimated $\delta^{15}\text{N}$ values. Box plots show 95% confidence intervals, quartiles, and mean for MCMC + ^{15}N . Diamonds show values determined by L2MN + ^{15}N . Dashed gray line shows “true” value from the forward model run. * indicates that the DIAZO model had only one dissolved inorganic nutrient pool, while the LIM models had two.

<https://doi.org/10.1371/journal.pone.0199123.g007>

instances, the L2MN + ^{15}N also did a reasonable job of recovering most $\delta^{15}\text{N}$ values. However, occasionally estimates generated by this approach were far from the “true” values, particularly when dealing with compartments through which the L2MN + ^{15}N approach predicted relatively little energy throughput. Because of these occasional large misfits, the r^2 value of a linear regression was lower (0.76).

Considerations for use with in situ data sets

In this study, we used simulated datasets from four configurations of dynamical (forward) models to test the efficacy with which static, mass-balanced LIM approaches recover ecosystem structure when only given inputs that are typically measured in situ. This approach has proven a powerful way to assess investigate LIM methodology [34, 59], although it is not without its limitations due to the inherent differences between dynamic simulations and steady-state LIM models [60]. To alleviate these issues, our model simulations were run to steady state with constant physical forcing (i.e. constant upwelling in NEMURO simulations; constant river input in DIAZO simulations). The ocean is seldom at a true steady state, suggesting that it could be fruitful for future studies to use results derived from non-steady state, three-dimensional, coupled biogeochemical models as inputs for assessing LIM accuracy. However, prior work has suggested that LIM results based on transient states are as accurate as those derived from true steady-state conditions [34].

When using LIM data assimilation techniques it is important to consider the inherent biases of the L2MN and MCMC approaches. Well known biases associated with the L2MN approach have been assessed in other manuscripts [27, 29, 30] and are related to the L2MN approach’s goal of minimizing total flow through the ecosystem. In our simulations, this was

apparent in the L2MN approach's attempt to minimize recycled production (minimizing NH_4^+ production by multiple compartments while maximizing N_2 fixation) and minimize the number of trophic steps through the zooplankton (thus typically underestimating mesozooplankton trophic level and overestimating secondary production). The MCMC approach, however, has subtle biases of its own that must be considered. Specifically, a greater portion of the solution space tends to exist at high total system throughput (i.e. sum of all flows) than at low total system throughput. Thus the MCMC mean solutions (which average across all possible solution sets) tend to be biased towards solutions that increase total flow, often by including many trophic steps and enhanced recycling.

In addition to these biases, LIM model results can be affected by the assumed ecosystem structure used to construct the model. For instance, increased aggregation of functional groups (i.e. inclusion of less compartments) was shown to decrease LIM model accuracy in a tidal system [61]. In pelagic systems, even splitting detritus into three size-structured detritus compartments can substantially impact the relative contribution of different phytoplankton groups to total export [4]. Unfortunately, the appropriate level of aggregation and the true ecosystem structure are seldom known a priori and must be estimated by the investigator. In this way, our decision to use different ecosystem structure for the LIM and the two forward models allows us to simulate the difficulties found in the field. For instance, the LIM model allows diazotrophy, which is absent from the NEMURO forward model. Thus the LIM model consistently overestimates the "true" rate of N_2 fixation (zero). An important result of our study was that the inclusion of ^{15}N data improved the LIM model's ability to reconstruct the "true" values despite these differences in model structure.

Regardless of which approach is used (MCMC or L2MN), the inclusion of ^{15}N data provides additional constraint on the system. This agrees with the finding of Vézina and Pahlow [34] that inverse approaches using multiple currencies (e.g. C and N, or N and ^{15}N) were more accurate than approaches using only a single currency. When applied to natural pelagic ecosystems that are usually highly underconstrained due to the difficulty of measuring planktonic rates in situ, we expect that both ^{15}N approaches will outperform the results of LIM models without this additional data source. However, although our forward models and LIM approaches assumed the same known isotopic fractionation factors, in situ fractionation factors should be assumed to have some uncertainty to them. Indeed, our understanding of taxonomic diversity in fractionation processes is still evolving. For instance, the isotopic fractionation coefficients for zooplankton (ϵ_{exc} and ϵ_{eg}) together control the trophic enrichment factor (TEF) of consumers, which has typically been assumed to be in the range of 3–3.5 for a diverse suite of organisms. However, recent evidence [62, 63] suggests that the TEF of protozoans is much lower (less than 1), thus necessitating lower values for ϵ_{exc} and ϵ_{eg} . If accurate 95% confidence intervals on $\delta^{15}\text{N}$ values are required, it is thus likely necessary to incorporate uncertainty in isotopic fractionation coefficients.

When used with in situ data sets, the MCMC+ ^{15}N approach can be easily adapted to incorporate variable stoichiometric data (with or without $\delta^{15}\text{N}$ data) if such datatypes are available. It would be trivial to replace ∂_K and ∂_U (the vectors containing the known and unknown/varying $\delta^{15}\text{N}$ values, respectively) with vectors including C:N, N:P, or $\delta^{13}\text{C}$ data, each of which could be used to update the approximate equations stored in matrix A. In this way, our approach is best seen as a flexible tool that can be used to assimilate diverse available datasets into a constrained estimate of ecosystem structure.

APPENDIX 1 –approximate equalities

Input measurements used for all LIM approaches

$$\text{NFixDTM} + \text{NO}_3 \rightarrow \text{DTM} + \text{NH}_4 \rightarrow \text{DTM} - \text{DTM} \rightarrow \text{DOM} + \text{NFixCYA} + \text{NO}_3 \rightarrow \text{CYA} + \text{NH}_4 \rightarrow \text{CYA} - \text{CYA} \rightarrow \text{DOM} = \text{NPP}$$

$$\text{DTM} \rightarrow \text{MES} = \text{MesozooGrazing}$$

$$\text{NO}_3 \rightarrow \text{DTM} + \text{NO}_3 \rightarrow \text{CYA} = \text{NitrateUptake}$$

$$\text{DET} \rightarrow \text{Sink} = \text{SedimentTrapExport}$$

¹⁵N Mass balance equations used for L2MN+¹⁵N and MCMC+¹⁵N

$$\begin{aligned} &R_{\text{upno}_3} \times \text{EXT} \rightarrow \text{NO}_3 - (R_{\text{NO}_3} \times \alpha_{\text{no}_3}) \times \text{NO}_3 \rightarrow \text{CYA} - (R_{\text{NO}_3} \times \alpha_{\text{no}_3}) \times \text{NO}_3 \rightarrow \text{DTM} = 0 \\ &- (R_{\text{nh}_4} \times \alpha_{\text{nh}_4}) \times \text{NH}_4 \rightarrow \text{DTM} - (R_{\text{nh}_4} \times \alpha_{\text{nh}_4}) \times \text{NH}_4 \rightarrow \text{CYA} + (R_{\text{hnf}} \times \alpha_{\text{exc}}) \times \text{HNF} \rightarrow \text{NH}_4 + \\ &(R_{\text{mic}} \times \alpha_{\text{exc}}) \times \text{MIC} \rightarrow \text{NH}_4 + (R_{\text{mes}} \times \alpha_{\text{exc}}) \times \text{MES} \rightarrow \text{NH}_4 + (R_{\text{DOM}} \times \alpha_{\text{sol}}) \times \text{DOM} \rightarrow \text{NH}_4 = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{nfix}} \times \text{NFixCYA} + (R_{\text{NO}_3} \times \alpha_{\text{no}_3}) \times \text{NO}_3 \rightarrow \text{CYA} + (R_{\text{nh}_4} \times \alpha_{\text{nh}_4}) \times \text{NH}_4 \rightarrow \text{CYA} - R_{\text{cya}} \times \text{CYA} \rightarrow \text{HNF} - \\ &R_{\text{cya}} \times \text{CYA} \rightarrow \text{MIC} - R_{\text{cya}} \times \text{CYA} \rightarrow \text{DET} - R_{\text{cya}} \times \text{CYA} \rightarrow \text{DOM} = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{nfix}} \times \text{NFixDTM} + (R_{\text{NO}_3} \times \alpha_{\text{no}_3}) \times \text{NO}_3 \rightarrow \text{DTM} + (R_{\text{nh}_4} \times \alpha_{\text{nh}_4}) \times \text{NH}_4 \rightarrow \text{DTM} - \\ &R_{\text{dtm}} \times \text{DTM} \rightarrow \text{MIC} - R_{\text{dtm}} \times \text{DTM} \rightarrow \text{MES} - R_{\text{dtm}} \times \text{DTM} \rightarrow \text{DET} - R_{\text{dtm}} \times \text{DTM} \rightarrow \text{DOM} = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{cya}} \times \text{CYA} \rightarrow \text{HNF} + R_{\text{det}} \times \text{DET} \rightarrow \text{HNF} - R_{\text{hnf}} \times \text{HNF} \rightarrow \text{MIC} - R_{\text{hnf}} \times \text{HNF} \rightarrow \text{MES} - \\ &(R_{\text{hnf}} \times \alpha_{\text{exc}}) \times \text{HNF} \rightarrow \text{NH}_4 - (R_{\text{hnf}} \times \alpha_{\text{eg}}) \times \text{HNF} \rightarrow \text{DET} - (R_{\text{hnf}} \times \alpha_{\text{exc}}) \times \text{HNF} \rightarrow \text{DOM} = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{cya}} \times \text{CYA} \rightarrow \text{MIC} + R_{\text{dtm}} \times \text{DTM} \rightarrow \text{MIC} + R_{\text{hnf}} \times \text{HNF} \rightarrow \text{MIC} + R_{\text{det}} \times \text{DET} \rightarrow \text{MIC} - R_{\text{mic}} \times \text{MIC} \rightarrow \text{MES} - \\ &(R_{\text{mic}} \times \alpha_{\text{exc}}) \times \text{MIC} \rightarrow \text{NH}_4 - (R_{\text{mic}} \times \alpha_{\text{eg}}) \times \text{MIC} \rightarrow \text{DET} - \\ &(R_{\text{mic}} \times \alpha_{\text{exc}}) \times \text{MIC} \rightarrow \text{DOM} = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{dtm}} \times \text{DTM} \rightarrow \text{MES} + R_{\text{hnf}} \times \text{HNF} \rightarrow \text{MES} + R_{\text{mic}} \times \text{MIC} \rightarrow \text{MES} + R_{\text{det}} \times \text{DET} \rightarrow \text{MES} - R_{\text{mes}} \times \text{MES} \rightarrow \text{HTL} - \\ &(R_{\text{mes}} \times \alpha_{\text{exc}}) \times \text{MES} \rightarrow \text{NH}_4 - (R_{\text{mes}} \times \alpha_{\text{eg}}) \times \text{MES} \rightarrow \text{DET} - \\ &(R_{\text{mes}} \times \alpha_{\text{exc}}) \times \text{MES} \rightarrow \text{DOM} = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{dtm}} \times \text{DTM} \rightarrow \text{DET} + R_{\text{cya}} \times \text{CYA} \rightarrow \text{DET} + (R_{\text{hnf}} \times \alpha_{\text{eg}}) \times \text{HNF} \rightarrow \text{DET} + (R_{\text{mic}} \times \alpha_{\text{eg}}) \times \text{MIC} \rightarrow \text{DET} + \\ &(R_{\text{mes}} \times \alpha_{\text{eg}}) \times \text{MES} \rightarrow \text{DET} - R_{\text{det}} \times \text{DET} \rightarrow \text{HNF} - R_{\text{det}} \times \text{DET} \rightarrow \text{MIC} - R_{\text{det}} \times \text{DET} \rightarrow \text{MES} - \\ &(R_{\text{det}} \times \alpha_{\text{sol}}) \times \text{DET} \rightarrow \text{DOM} - R_{\text{det}} \times \text{DET} \rightarrow \text{Sink} = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{dtm}} \times \text{DTM} \rightarrow \text{DOM} + R_{\text{cya}} \times \text{CYA} \rightarrow \text{DOM} + (R_{\text{hnf}} \times \alpha_{\text{exc}}) \times \text{HNF} \rightarrow \text{DOM} + (R_{\text{mic}} \times \alpha_{\text{exc}}) \times \text{MIC} \rightarrow \text{DOM} + \\ &(R_{\text{mes}} \times \alpha_{\text{exc}}) \times \text{MES} \rightarrow \text{DOM} + (R_{\text{det}} \times \alpha_{\text{sol}}) \times \text{DET} \rightarrow \text{DOM} - \\ &(R_{\text{dom}} \times \alpha_{\text{sol}}) \times \text{DOM} \rightarrow \text{NH}_4 = 0 \end{aligned}$$

$$\begin{aligned} &R_{\text{upno}_3} \times \text{EXT} \rightarrow \text{NO}_3 + R_{\text{nfix}} \times \text{NFixCYA} + R_{\text{nfix}} \times \text{NFixDTM} - R_{\text{det}} \times \text{DET} \rightarrow \text{Sink} - \\ &R_{\text{mes}} \times \text{MES} \rightarrow \text{HTL} = 0 \end{aligned}$$

In all equations R_x refers to the ¹⁵N:¹⁴N isotopic ratio of compartment x, which is computed from $\delta^{15}\text{N}$ values using the equation $R_x = \delta^{15}\text{N}_x \times R_{\text{N}_2} / 1000 + R_{\text{N}_2}$, where R_{N_2} is the ¹⁵N:¹⁴N isotopic ratio of atmospheric dinitrogen gas. In all equations α_y refers to the isotopic fractionation coefficient for process y and is calculated from the isotopic fractionation factor (ϵ_y) for process y according to the equation $\alpha_y = \exp(\epsilon_y/1000)$. Fractionation factors used in this study were taken from Yoshikawa et al. [41] and had values of $\epsilon_{\text{NO}_3} = -5\%$, $\epsilon_{\text{NH}_4} = -10\%$, $\epsilon_{\text{exc}} = -5\%$, $\epsilon_{\text{eg}} = -2\%$, $\epsilon_{\text{rem}} = -1\%$.

Supporting information

S1 Supplementary Text. A detailed description of our implementation of the MCMC+¹⁵N LIM method.

(PDF)

S1 Fig. Histograms of nitrate uptake fraction (x-axis) and nitrogen fixation fraction (y-axis) and a scatter plot of nitrogen fixation fraction against nitrate uptake fraction as computed by the MCMC+¹⁵N approach. For comparison, the MCMC approach mean value is shown in cyan diamond (with 95% confidence interval) and L2MN and L2MN+¹⁵N values are shown in pink and dark red, respectively.

(JPG)

S2 Fig. Histograms of mesozooplankton secondary production (x-axis) and mesozooplankton trophic level (y-axis) and a scatter plot of trophic level against secondary production as computed by the MCMC+ ^{15}N approach. For comparison, the MCMC approach mean value is shown in cyan diamond (with 95% confidence interval) and L2MN and L2MN+ ^{15}N values are shown in pink and dark red, respectively.
(JPG)

Acknowledgments

We are grateful to our colleagues in the CCE LTER Program and NOAA Restore Act Bluefin Tuna Project for many discussions about ways to synthesize cruise data. The code used in this manuscript (and an example setup file) can be downloaded from GitHub at: <https://github.com/stukel-lab/N15-LIM>.

Author Contributions

Conceptualization: Michael R. Stukel, Moira Décima.

Investigation: Michael R. Stukel, Moira Décima, Thomas B. Kelly.

Methodology: Thomas B. Kelly.

Writing – original draft: Michael R. Stukel.

Writing – review & editing: Michael R. Stukel, Moira Décima, Thomas B. Kelly.

References

1. Christensen V, Pauly D. ECOPATH II: a software for balancing steady-state ecosystem models and calculating network characteristics. *Ecological Modelling*. 1992; 61(3–4):169–85. ISI:A1992HZ63300002.
2. van Oevelen D, Van den Meersche K, Meysman FJR, Soetaert K, Middelburg JJ, Vézina AF. Quantifying food web flows using linear inverse models. *Ecosystems*. 2010; 13(1):32–45. <https://doi.org/10.1007/s10021-009-9297-6> ISI:000275172700002.
3. Vézina AF, Platt T. Food web dynamics in the ocean .1. Best estimates of flow networks using inverse methods. *Mar Ecol-Prog Ser*. 1988; 42(3):269–87. ISI:A1988M287000007.
4. Stukel MR, Landry MR. Contribution of picophytoplankton to carbon export in the equatorial Pacific: A re-assessment of food-web flux inferences from inverse models. *Limnology and Oceanography*. 2010; 55(6):2669–85.
5. Rau GH, Teyssie JL, Rassoulzadegan F, Fowler SW. $^{13}\text{C}/^{12}\text{C}$ and $^{15}\text{N}/^{14}\text{N}$ variations among size-fractionated marine particles: implications for their origin and trophic relationships. *Mar Ecol-Prog Ser*. 1990; 59(1–2):33–8. <https://doi.org/10.3354/meps059033> WOS:A1990CJ40300003.
6. Fry B, Sherr EB. $\delta^{13}\text{C}$ measurements as indicators of carbon flow in marine and freshwater ecosystems. *Stable isotopes in ecological research*: Springer; 1989. p. 196–229.
7. Post DM. Using stable isotopes to estimate trophic position: models, methods, and assumptions. *Ecology*. 2002; 83(3):703–18.
8. Fry B, Quiñones RB. Biomass spectra and stable isotope indicators of trophic level in zooplankton of the northwest Atlantic. *Mar Ecol-Prog Ser*. 1994; 112:201–4.
9. Fawcett SE, Lomas M, Casey JR, Ward BB, Sigman DM. Assimilation of upwelled nitrate by small eukaryotes in the Sargasso Sea. *Nature Geoscience*. 2011; 4(10):717–22. <https://doi.org/10.1038/ngeo1265> WOS:000295403900021.
10. Lassalle G, Chouvelon T, Bustamante P, Niquil N. An assessment of the trophic structure of the Bay of Biscay continental shelf food web: Comparing estimates derived from an ecosystem model and isotopic data. *Prog Oceanogr*. 2014; 120:205–15. <https://doi.org/10.1016/j.pcean.2013.09.002>
11. Eldridge PM, Cifuentes LA, Kaldy JE. Development of a stable-isotope constraint system for estuarine food-web models. *Mar Ecol-Prog Ser*. 2005; 303:73–90. <https://doi.org/10.3354/meps303073> WOS:000234214500006.

12. van Oevelen D, Soetaert K, Middelburg JJ, Herman PMJ, Moodley L, Hamels I, et al. Carbon flows through a benthic food web: Integrating biomass, isotope and tracer data. *J Mar Res.* 2006; 64(3):453–82. <https://doi.org/10.1357/002224006778189581> WOS:000240251400006.
13. Phillips DL, Inger R, Bearhop S, Jackson AL, Moore JW, Parnell AC, et al. Best practices for use of stable isotope mixing models in food-web studies. *Can J Zool.* 2014; 92(10):823–35. <https://doi.org/10.1139/cjz-2014-0127> WOS:000345031300001.
14. Casciotti KL. Nitrogen and Oxygen Isotopic Studies of the Marine Nitrogen Cycle. *Annual Review of Marine Science.* 2016; 8(1):379–407. <https://doi.org/10.1146/annurev-marine-010213-135052> PMID: 26747521.
15. Knapp AN, Casciotti KL, Berelson WM, Prokopenko MG, Capone DG. Low rates of nitrogen fixation in eastern tropical South Pacific surface waters. *Proceedings of the National Academy of Sciences.* 2016; 113(16):4398–403. <https://doi.org/10.1073/pnas.1515641113> PMID: 26976587
16. Dore JE, Brum Jr., Tupas LM, Karl DM. Seasonal and interannual variability in sources of nitrogen supporting export in the oligotrophic subtropical North Pacific Ocean. *Limnology and Oceanography.* 2002; 47(6):1595–607. WOS:000179650200003.
17. White AE, Foster RA, Benitez-Nelson CR, Masque P, Verdeny E, Popp BN, et al. Nitrogen fixation in the Gulf of California and the Eastern Tropical North Pacific. *Prog Oceanogr.* 2013; 109:1–17. <https://doi.org/10.1016/j.pocean.2012.09.002> WOS:000315059200001.
18. Cabana G, Rasmussen JB. Comparison of aquatic food chains using nitrogen isotopes. *Proceedings of the National Academy of Sciences.* 1996; 93(20):10844–7.
19. Vander Zanden MJ, Cabana G, Rasmussen JB. Comparing trophic position of freshwater fish calculated using stable nitrogen isotope ratios ($\delta^{15}\text{N}$) and literature dietary data. *Canadian Journal of Fisheries and Aquatic Sciences.* 1997; 54(5):1142–58.
20. Sarà G, Sarà R. Feeding habits and trophic levels of bluefin tuna *Thunnus thynnus* of different size classes in the Mediterranean Sea. *Journal of Applied Ichthyology.* 2007; 23(2):122–7.
21. Stergiou KI, Karpouzi VS. Feeding habits and trophic levels of Mediterranean fish. *Reviews in Fish Biology and Fisheries.* 2002; 11(3):217–54.
22. Parnell AC, Phillips DL, Bearhop S, Semmens BX, Ward EJ, Moore JW, et al. Bayesian stable isotope mixing models. *Environmetrics.* 2013; 24(6):387–99. <https://doi.org/10.1002/env.2221>
23. Moore JW, Semmens BX. Incorporating uncertainty and prior information into stable isotope mixing models. *Ecology Letters.* 2008; 11(5):470–80. <https://doi.org/10.1111/j.1461-0248.2008.01163.x> PMID: 18294213
24. Stock BC, Semmens BX. Unifying error structures in commonly used biotracer mixing models. *Ecology.* 2016; 97(10):2562–9. <https://doi.org/10.1002/ecy.1517> PMID: 27859126
25. Kadoya T, Osada Y, Takimoto G. IsoWeb: a Bayesian isotope mixing model for diet analysis of the whole food web. *PloS one.* 2012; 7(7):e41057. <https://doi.org/10.1371/journal.pone.0041057> PMID: 22848427
26. Pacella SR, Lebreton B, Richard P, Phillips D, DeWitt TH, Niquil N. Incorporation of diet information derived from Bayesian stable isotope mixing models into mass-balanced marine ecosystem models: A case study from the Marennes-Oleron Estuary, France. *ecological modelling.* 2013; 267:127–37. <https://doi.org/10.1016/j.ecolmodel.2013.07.018> WOS:000324656600012.
27. Niquil N, Jackson GA, Legendre L, Delesalle B. Inverse model analysis of the planktonic food web of Takapoto Atoll (French Polynesia). *Mar Ecol-Prog Ser.* 1998; 165:17–29. <https://doi.org/10.3354/meps165017> WOS:000073828700002.
28. Kones JK, Soetaert K, van Oevelen D, Owino JO. Are network indices robust indicators of food web functioning? A Monte Carlo approach. *Ecological Modelling.* 2009; 220(3):370–82. <https://doi.org/10.1016/j.ecolmodel.2008.10.012> ISI:000263209400009.
29. Kones JK, Soetaert K, van Oevelen D, Owino JO, Mavuti K. Gaining insight into food webs reconstructed by the inverse method. *J Mar Syst.* 2006; 60(1–2):153–66. <https://doi.org/10.1016/j.jmarsys.2005.12.002> ISI:000236958900011.
30. Stukel MR, Landry MR, Ohman MD, Goericke R, Samo T, Benitez-Nelson CR. Do inverse ecosystem models accurately reconstruct plankton trophic flows? Comparing two solution methods using field data from the California Current. *J Mar Syst.* 2012; 91(1):20–33. <https://doi.org/10.1016/j.jmarsys.2011.09.004> WOS:000297891200003.
31. van Oevelen D, Duineveld G, Lavaleye M, Mienis F, Soetaert K, Heip CHR. The cold-water coral community as hotspot of carbon cycling on continental margins: A food-web analysis from Rockall Bank (northeast Atlantic). *Limnology and Oceanography.* 2009; 54(6):1829–44. <https://doi.org/10.4319/lo.2009.54.6.1829>

32. van Oevelen D, Soetaert K, Heip C. Carbon flows in the benthic food web of the Porcupine Abyssal Plain: The (un) importance of labile detritus in supporting microbial and faunal carbon demands. *Limnology and Oceanography*. 2012; 57(2):645–64.
33. Gontikaki E, van Oevelen D, Saetaert K, Witte U. Food web flows through a sub-arctic deep-sea benthic community. *Prog Oceanogr*. 2011. <https://doi.org/10.1016/j.pocean.2010.12.014>
34. Vézina AF, Pahlow M. Reconstruction of ecosystem flows using inverse methods: how well do they work? *J Mar Syst*. 2003; 40:55–77. [https://doi.org/10.1016/s0924-7963\(03\)00013-7](https://doi.org/10.1016/s0924-7963(03)00013-7) ISI:000182625300004.
35. Kishi MJ, Ito S-i, Megrey BA, Rose KA, Werner FE. A review of the NEMURO and NEMURO.FISH models and their application to marine ecosystem investigations. *Journal of oceanography*. 2011; 67(1):3–16.
36. Kishi MJ, Kashiwai M, Ware DM, Megrey BA, Eslinger DL, Werner FE, et al. NEMURO—a lower trophic level model for the North Pacific marine ecosystem. *Ecological Modelling*. 2007; 202(1–2):12–25. <https://doi.org/10.1016/j.ecolmodel.2006.08.021> ISI:000245311000004.
37. Li QP, Franks PJS, Landry MR, Goericke R, Taylor AG. Modeling phytoplankton growth rates and chlorophyll to carbon ratios in California coastal and pelagic ecosystems. *Journal of Geophysical Research-Biogeosciences*. 2010; 115:G04003. doi: 1029/2009JG001111 ISI:000282768500001.
38. Stukel MR, Coles VJ, Brooks MT, Hood RR. Top-down, bottom-up and physical controls on diatom-diazotroph assemblage growth in the Amazon River plume. *Biogeosciences*. 2014; 11:3259–78. <https://doi.org/10.5194/bg-11-3259-2014>
39. Subramaniam A, Yager PL, Carpenter EJ, Mahaffey C, Bjorkman K, Cooley S, et al. Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean. *Proc Natl Acad Sci U S A*. 2008; 105(30):10460–5. <https://doi.org/10.1073/pnas.0710279105> ISI:000258211600035. PMID: 18647838
40. Goes JI, Gomes HDR, Chekalyuk AM, Carpenter EJ, Montoya JP, Coles VJ, et al. Influence of the Amazon River discharge on the biogeography of phytoplankton communities in the western tropical north Atlantic. *Prog Oceanogr*. 2014; 120:29–40. <https://doi.org/10.1016/j.pocean.2013.07.010> WOS:000331019300002.
41. Yoshikawa C, Yamanaka Y, Nakatsuka T. An Ecosystem Model Including Nitrogen Isotopes: Perspectives on a Study of the Marine Nitrogen Cycle. *Journal of oceanography*. 2005; 61(5):921–42. <https://doi.org/10.1007/s10872-006-0010-5>
42. Landry MR, Ohman MD, Goericke R, Stukel MR, Tsyrklevich K. Lagrangian studies of phytoplankton growth and grazing relationships in a coastal upwelling ecosystem off Southern California. *Prog Oceanogr*. 2009; 83:208–16. <https://doi.org/10.1016/j.pocean.2009.07.026>
43. Stukel MR, Kahru M, Benitez-Nelson CR, Decima M, Goericke R, Landry MR, et al. Using Lagrangian-based process studies to test satellite algorithms of vertical carbon flux in the eastern North Pacific Ocean. *Journal of Geophysical Research: Oceans*. 2015; 120:7208–22. <https://doi.org/10.1002/2015JC011264>
44. Goericke R. The size structure of marine phytoplankton—What are the rules? *Calif Coop Ocean Fish Invest Rep*. 2011; 52:198–204. WOS:000299855500013.
45. Conroy BJ, Steinberg DK, Stukel MR, Goes JI, Coles VJ. Meso- and microzooplankton grazing in the Amazon River plume and western tropical North Atlantic. *Limnology and Oceanography*. 2016; 61(3):825–40. <https://doi.org/10.1002/lno.10261>
46. Weber SC, Carpenter EJ, Coles VJ, Yager PL, Goes J, Montoya JP. Amazon River influence on nitrogen fixation and export production in the western tropical North Atlantic. *Limnology and Oceanography*. 2016:n/a-n/a. <https://doi.org/10.1002/lno.10448>
47. Jackson GA, Eldridge PM. Food web analysis of a planktonic system off Southern California. *Prog Oceanogr*. 1992; 30(1–4):223–51. ISI:A1992JY67100007.
48. Richardson TL, Jackson GA, Ducklow HW, Roman MR. Carbon fluxes through food webs of the eastern equatorial Pacific: an inverse approach. *Deep-Sea Res I*. 2004; 51(9):1245–74. <https://doi.org/10.1016/j.dsr.2004.05.005> ISI:000223496500008.
49. Vézina AF. Construction of flow networks using inverse methods. In: Wulff F, Field JG, Mann KH, editors. *Network Analysis in Marine Ecology: Methods and Applications*. Berlin: Springer Verlag; 1989. p. 62–81.
50. Soetaert K, Van den Meersche K, Van Oevelen D. Package 'limSolve'. 2015.
51. Van den Meersche K, Soetaert K, Van Oevelen D. xSample(): An R function for sampling linear inverse problems. *Journal of Statistical Software, Code Snippets*. 2009; 30(1):1–15.
52. Saint-Béat B, Vézina AF, Asmus R, Asmus H, Niquil N. The mean function provides robustness to linear inverse modelling flow estimation in food webs: A comparison of functions derived from statistics and

- ecological theories. *Ecological Modelling*. 2013; 258:53–64. <http://dx.doi.org/10.1016/j.ecolmodel.2013.01.023>.
53. Altabet MA. Variations in Nitrogen Isotopic Composition between Sinking and Suspended Particles—Implications for Nitrogen Cycling and Particle Transformation in the Open Ocean. *Deep-Sea Research*. 1988; 35(4):535–54. ISI:A1988N960100005.
 54. Knapp AN, Sigman DM, Lipschultz F. N isotopic composition of dissolved organic nitrogen and nitrate at the Bermuda Atlantic time-series study site. *Glob Biogeochem Cycle*. 2005;19(1). <https://doi.org/10.1029/2004gb002320> WOS:000227874600002.
 55. Sigman DM, Casciotti KL, Andreani M, Barford C, Galanter M, Bohlke JK. A bacterial method for the nitrogen isotopic analysis of nitrate in seawater and freshwater. *Analytical Chemistry*. 2001; 73(17):4145–53. <https://doi.org/10.1021/ac010088e> WOS:000170859300024. PMID: 11569803
 56. Checkley DM, Entzeroth LC. Elemental and isotopic fractionation of carbon and nitrogen by marine, planktonic copepods and implications to the marine nitrogen cycle. *J Plankton Res*. 1985; 7(4):553–68. ISI:A1985AMS8100010.
 57. Soetaert K, Van den Meersche K, van Oevelen D. *limSolve: Solving linear inverse models*. R package version2009.
 58. Sailley SF, Ducklow HW, Moeller HV, Fraser WR, Schofield OM, Steinberg DK, et al. Carbon fluxes and pelagic ecosystem dynamics near two western Antarctic Peninsula Adelie penguin colonies: an inverse model approach. *Mar Ecol-Prog Ser*. 2013; 492:253–72. <https://doi.org/10.3354/meps10534> WOS:000326439800019.
 59. Vézina AF, Berreville F, Loza S. Inverse reconstructions of ecosystem flows in investigating regime shifts: impact of the choice of objective function. *Prog Oceanogr*. 2004; 60(2–4):321–41. <https://doi.org/10.1016/j.pocean.2004.02.012> ISI:000221508700011.
 60. Steele JH. Assessment of some linear food web methods. *J Mar Syst*. 2009;76(1–2):186–94. <https://doi.org/10.1016/j.jmarsys.2008.05.012> ISI:000263851000014.
 61. Johnson GA, Niquil N, Asmus H, Bacher C, Asmus R, Baird D. The effects of aggregation on the performance of the inverse method and indicators of network analysis. *Ecological Modelling*. 2009; 220(23):3448–64. <https://doi.org/10.1016/j.ecolmodel.2009.08.003> ISI:000272371300020.
 62. Décima M, Landry MR, Bradley CJ, Fogel ML. Alanine $\delta^{15}\text{N}$ trophic fractionation in heterotrophic protists. *Limnology and Oceanography*. 2017. <https://doi.org/10.1002/lno.10567>
 63. Gutierrez-Rodriguez A, Decima M, Popp BN, Landry MR. Isotopic invisibility of protozoan trophic steps in marine food webs. *Limnology and Oceanography*. 2014; 59(5):1590–8. <https://doi.org/10.4319/lo.2014.59.5.1590> WOS:000345462100011.