Do inverse ecosystem models accurately reconstruct plankton trophic flows? Comparing two solution methods using field data from the California Current

Michael R. Stukel [a,*], Michael R. Landry [a], Mark D. Ohman [a], Ralf Goericke [a], Ty Samo [a], Claudia R. Benitez-Nelson [b]

[a] *Scripps Institution of Oceanography, University of California at San Diego, 9500 Gilman Dr., La Jolla, CA 92037, USA*
[b] *University of South Carolina, Columbia, SC 29208, USA*

ARTICLE INFO

ABSTRACT

Despite the increasing use of linear inverse modeling techniques to elucidate fluxes in undersampled marine ecosystems, the accuracy with which they estimate food web flows has not been resolved. New Markov Chain Monte Carlo (MCMC) solution methods have also called into question the biases of the commonly used L_2 minimum norm (L_2MN) solution technique. Here, we test the abilities of MCMC and L_2MN methods to recover field-measured ecosystem rates that are sequentially excluded from the model input. For data, we use experimental measurements from process cruises of the California Current Ecosystem (CCE-LTER) Program that include rate estimates of phytoplankton and bacterial production, micro- and mesozooplankton grazing, and carbon export from eight study sites varying from rich coastal upwelling to offshore oligotrophic conditions. Both the MCMC and L_2MN methods predicted well-constrained rates of protozoan and mesozooplankton grazing with reasonable accuracy, but the MCMC method overestimated primary production. The MCMC method more accurately predicted the poorly constrained rate of vertical carbon export than the L_2MN method, which consistently overestimated export. Results involving DOC and bacterial production were equivocal. Overall, when primary production is provided as model input, the MCMC method gives a robust depiction of ecosystem processes. Uncertainty in inverse ecosystem models is large and arises primarily from solution under-determinacy. We thus suggest that experimental programs focusing on food web fluxes expand the range of experimental measurements to include the nature and fate of detrital pools, which play large roles in the model.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The ability to compare ecosystem states quantitatively is essential for elucidating the mechanisms underlying marine trophic interactions and their relationships to biogeochemistry. Unfortunately, reconstruction of food webs is hampered generally by the sparsity of ecological rate measurements. Linear inverse models (LIM) are powerful data assimilation tools that have been developed to address this problem of ecosystem underdeterminacy. LIM are designed to integrate biomass assessments, rate measurements, and a priori knowledge of trophic structure and organismal capabilities into best estimates of the flow of energy or nutrient flows through an ecosystem. Despite data limitations, pioneering work in inverse modeling of plankton ecosystems (Jackson and Eldridge, 1992; Vézina and Platt, 1988) has allowed parameter estimation and comparisons in a variety of marine habitats. More recently, LIM have been used for diverse purposes including comparative analysis of bloom ecosystems (Daniels et al., 2006), integration of rate measurements with stable isotope determinations of trophic position in benthic communities (van Oevelen et al., 2006), and to compare responses to nutrient enrichment in different coastal systems (Olsen et al., 2006). Nonetheless, the lack of objective methods for assessing system constraints (analogous to model errors) has largely confounded efforts to evaluate the accuracy of model solutions for unmeasured rates.

LIM are based on a system of equations $A\mathbf{x} = b$ that quantify the mass balance and rate measurements that constrain the ecosystem (van Oevelen et al., 2010; Vézina and Platt, 1988). They also incorporate a series of inequality constraints $G\mathbf{x} \geq h$, which represent known limits on the biology of ecosystem components (for instance gross growth efficiency). However, the paucity of ecological measurements relative to modeled flows (variables) leads invariably to an under-determined system, and hence to infinite solutions that can fit sparse data. To choose from among these solutions, investigators have traditionally used the L_2 minimum norm (L_2MN) approach,

which selects the solution vector that minimizes the sum of squared flow values. While this approach is parsimonious in a mathematical sense (Niquil et al., 1998; Vézina and Platt, 1988), it imposes an artificial structure on the emergent ecosystem depiction. In particular, to achieve minimum flows, the L_2MN consistently inflates the respiration of consumers at lower trophic levels (Stukel and Landry, 2010), minimizes the trophic level of higher consumers (Niquil et al., 1998), and selects values for flows that lie on the outer bounds of the allowable solution space (Steele, 2009; Stukel and Landry, 2010).

The Markov Chain Monte Carlo (MCMC) technique is an alternative approach for solving the under-constrained inverse ecosystem problem by using random walk techniques to fully sample the solution space and computing maximum likelihood results (Kones et al., 2006; Kones et al., 2009; Soetaert and van Oevelen, 2009; van Oevelen et al., 2010). Unlike the L_2MN, this method makes no assumption about underlying ecosystem structure, but chooses as most likely the mean value of any given flow from among the set of all solutions that fit the available data. It thus determines solution statistics for each variable of interest (Kones et al., 2009) while avoiding the L_2MN's tendency to choose extreme values (Stukel and Landry, 2010).

While the complete sampling of the solution space by the MCMC offers advantages relative to the L_2MN, no studies have shown that it provides better approximations of ecosystem fluxes. To date, most studies addressing the efficacy of inverse techniques have used simulated ecosystems in which the ecologically relevant components were prescribed (Vézina and Pahlow, 2003; Vézina et al., 2004). However, simulated ecosystems behave according to their constructs in easily predictable ways, while the underlying structures of actual ecosystems are largely unknown, and hence less likely to be described accurately by a model.

In the present study, we compare alternative L_2MN and MCMC solutions of inverse ecosystem models based on data from two cruises of the California Current Ecosystem, Long-Term Ecological Research (CCE-LTER) Program. During CCE cruises in May 2006 (P0605) and April 2007 (P0704), a total of eight water parcels were tracked for an average of 4 days each while various processes – including phytoplankton growth rate, protozoan and mesozooplankton grazing, net phytoplankton accumulation, vertical carbon flux and bacterial growth rates – were measured experimentally. By running the model with both L_2MN and the MCMC solution methods while sequentially withholding measurements, we are able to assess and compare each minimization scheme's ability to predict the measured data. We find that the L_2MN method more accurately estimates phytoplankton production (when it is not a measured constraint). In contrast, if primary production is a model input the MCMC method was a better predictor of ecosystem flows.

2. Methods

2.1. Model structure

Structuring an inverse ecosystem model involves inevitable trade-offs between constraints (fewer compartments, less complexity) and broad depiction of ecosystem processes (more compartments, greater complexity). The level of complexity is a matter of judgment, but has been shown to influence model results (Stukel and Landry, 2010). Since the primary goal of this study was to compare methodologies, we chose to err on the side of greater constraint, hence fewer compartments. We thus constructed a simple ecosystem (Fig. 1) with one phytoplankton group (Phy), three size-structured grazing groups (Hnf, Mic, Mes), bacteria (Bac), detritus (Det), and dissolved organic carbon (DOC). Each grazing group was allowed to feed upon phytoplankton and smaller consumers. Bacterivory was allowed only for nanoflagellates (Hnf) and microzooplankton (Mic). Egestion was incorporated as a flux from grazers to detritus, while phytoplankton
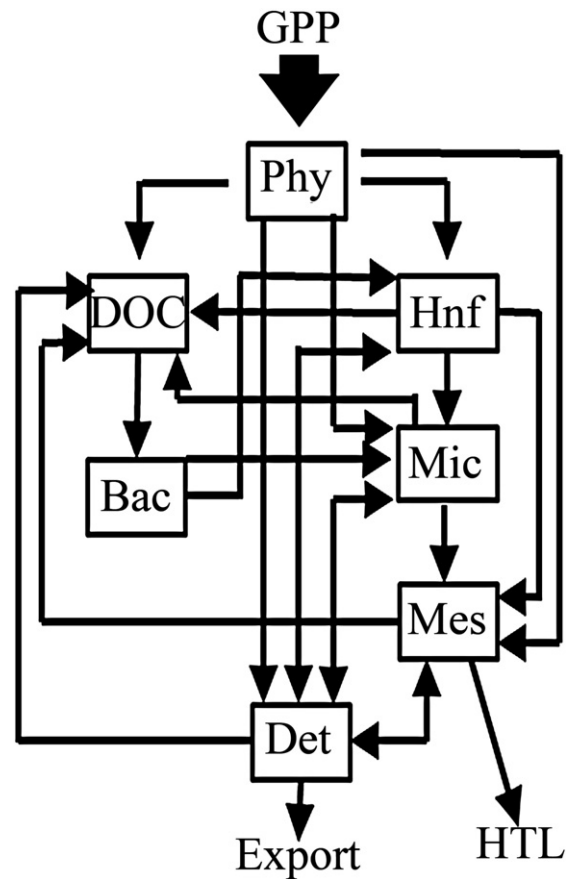
Fig. 1. Model structure. Model compartments are phytoplankton (Phy), heterotrophic nanoflagellates (Hnf), microzooplankton (Mic), mesozooplankton (Mes), bacteria (Bac), detritus (Det), and dissolved organic carbon (DOC). GPP = gross primary production. Arrows show the direction of flow through the ecosystem. The two export terms in the model (Det to export and Mes to HTL) are interpreted as sinking particulate organic carbon and mesozooplankton loss to higher trophic levels, respectively. Not shown in the diagram are respiratory losses of each of the 5 living compartments.

contributed to detritus through cell death. DOC was produced by direct exudation from phytoplankton and the excretion and sloppy feeding of grazers. DOC uptake by bacteria was modeled as a net uptake to avoid the unbounded flows that result from inclusion of DOC production by bacteria, a process that is poorly constrained by field and laboratory measurements. Energy was dissipated through the respiration of each group, by the vertical flux of detritus sinking out of the euphotic zone, and by higher trophic level consumption of mesozooplankton. The model thus had a total of 29 flows (Fig. 1), constrained by experimental measurements of net primary production, bacterial production, herbivorous grazing by the protistan community, herbivorous grazing by mesozooplankton (Mes), and export of sinking particulates (some measurements were not made for all of the eight experiments). The addition of seven mass-balance constraints (discussed below) led to a maximum of 12 constraints on the ecosystem model.

2.2. Ecological measurements

Model data (Table 1) are from CCE-LTER process cruises in May 2006 (P0605) and April 2007 (P0704). During these cruises, homogeneous water parcels were located using site surveys with a Moving Vessel Profiler (MVP: ODIM Brooke Ocean; Ohman, unpub.) and marked with a drift array drogued at 15 m in the surface mixed layer (Landry et al., 2009). Water parcels were typically followed for

Table 1

Measured inputs to the model. Distance is initial position of drifter from shore. μ and m are the phytoplankton specific growth rate and protozoan grazing rates measured by the dilution technique. MesoGr is the mesozooplankton grazing rate as determined by the gut pigment technique. POCFlux is vertical carbon export. BacProd is bacterial production. ΔChl is the *in situ* rate of change of chlorophyll over the course of the cycle. Bio denotes biomass. Euphotic depth is the depth of integration for biological measurements. Mean ± 95% C.L. ND = no data. Column headings indicate cruise (e.g., 0605) and cycle number (e.g., 1).

		0605-1	0605-2	0605-3	0605-4	0605-5	0704-1	0704-2	0704-3
Distance	Km	29	166	11	89	356	39	255	63
^{14}C-PP	mg C m^{-2} d^{-1}	4184 ± 1767	563 ± 28	4382 ± 412	1474 ± 251	483 ± 159	1233 ± 815	587 ± 90	2314 ± 910
μ	d^{-1}	0.273 ± 0.060	0.159 ± 0.051	0.243 ± 0.168	0.346 ± 0.068	0.073 ± 0.053	0.293 ± 0.139	0.136 ± 0.032	0.285 ± 0.030
m	d^{-1}	0.125 ± 0.054	0.123 ± 0.063	0.175 ± 0.112	0.154 ± 0.053	0.068 ± 0.042	0.157 ± 0.081	0.166 ± 0.043	0.158 ± 0.048
MesoGr	d^{-1}	0.088 ± 0.046	0.057 ± 0.029	0.112 ± 0.092	0.065 ± 0.007	0.013 ± 0.003	0.475 ± 0.089	0.044 ± 0.008	0.219 ± 0.133
POCFlux	mg C m^{-2} d^{-1}	322 ± 200	72 ± 44	ND	132 ± 49	76 ± 74	83 ± 14	37 ± 11	129 ± 61
BacProd	mg C m^{-2} d^{-1}	ND	ND	ND	ND	ND	53 ± 25	22 ± 12	33 ± 15
ΔChl	d^{-1}	0.070 ± 0.238	−0.004 ± 0.109	0.002 ± 0.589	0.142 ± 0.154	−0.173 ± 0.152	−0.253 ± 0.164	0.014 ± 0.199	−0.069 ± 0.519
PhyBio	mg C m^{-2}	7451 ± 2094	2262 ± 676	7333 ± 2024	3085 ± 629	1653 ± 323	2613 ± 1019	2079 ± 165	2239 ± 333
HnfBio	mg C m^{-2}	631 ± 42	575 ± 32	397 ± 48	406 ± 25	298 ± 11	410 ± 20	457 ± 9	396 ± 16
MicBio[a]	mg C m^{-2}	1733 ± 1034	128 ± 128	1390 ± 2082	99 ± 228	243 ± 69	262 ± 106	41 ± 48	203 ± 265
MesBio	mg C m^{-2}	1335 ± 532	365 ± 123	1329 ± 408	757 ± 188	243 ± 44	2695 ± 1077	391 ± 87	1715 ± 393
BacBio	mg C m^{-2}	756 ± 263	810 ± 29	489 ± 77	520 ± 45	917 ± 53	539 ± 49	907 ± 54	885 ± 93
POC	mg C m^{-2}	12,146 ± 3331	4637 ± 210	10,523 ± 3155	5320 ± 694	4845 ± 669	4618 ± 1416	4839 ± 174	6022 ± 752
DOC	mg C m^{-2}	20,801 ± 7261	69,212 ± 8569	9023 ± 3958	30,761 ± 1460	54,105 ± 4014	30,231 ± 2228	112,472 ± 104,020	52,395 ± 2052
MaxHNFBio	mg C m^{-2}	694	850	554	508	359	463	612	458
MaxMicBio[a]	mg C m^{-2}	5718	1648	10,483	1376	747	2194	653	1259
MaxMesBio	mg C m^{-2}	2120	549	1725	890	280	4095	488	1942
MaxPhyDensity	mg C m^{-3}	431	90.76	822.34	95.15	65.6	116.23	54.66	59.29
MaxHNFDensity	mg C m^{-3}	20	17.82	33.8	21.14	7.14	14.51	10.54	8.85
MaxMicDensity[a]	mg C m^{-3}	313	53.3	724.26	44.44	23.94	57.94	39.42	15.59
MaxBacDensity	mg C m^{-3}	29	11	30	13	20	14	14	16
Temp	°C	11.7 ± 0.4	13.8 ± 0.2	12.9 ± 0.3	14.2 ± 0.2	15.0 ± 0.2	11.9 ± 0.2	13.9 ± 0.1	11.8 ± 0.3
Euphotic depth	M	50	100	25	50	90	50	90	85

[a] Maximum microzooplankton concentration terms include autotrophic dinoflagellates as they are likely phagotrophic.

a 4-day experimental "cycle" while ecological rates were measured simultaneously with net changes in the euphotic zone phytoplankton community. Landry et al. (2009) showed that net changes in the phytoplankton community could be predicted to first order by the differences between measured rates of phytoplankton growth and combined grazing pressures of herbivorous proto- and metazoans, thus suggesting that the experimental measurements accurately reflected the processes determining ecosystem changes.

Phytoplankton carbon production was determined by ^{14}C uptake (^{14}C-PP) measurements made from 4-L samples incubated *in situ* at 6 depths on our drift array and subsampled (250 mL) in triplicate for decay counts, with a separate 250-mL dark bottle used to correct for non-photosynthetic carbon uptake or adsorption onto particles. Rate estimates for phytoplankton growth and grazing were made as chlorophyll-based euphotic-zone averages from the dilution method for growth and microzooplankton grazing, and from the gut pigments of mesozooplankton collected in day–night net tows (Landry et al., 2009 and Ohman, unpubl. data). Microzooplankton grazing rates are scaled to carbon equivalents by multiplying vertically integrated ^{14}C-PP estimates times the ratios of protozoan grazing to phytoplankton specific growth rate from dilution results. Mesozooplankton grazing rates are similarly scaled to ^{14}C-PP carbon units by their initial chlorophyll-based ratios of grazing to growth. Net rates of change in phytoplankton community biomass were assessed from daily changes in vertically integrated chlorophyll concentrations at the drifter location multiplied by the mean ratio of chlorophyll to phytoplankton carbon biomass calculated for each cycle from epifluorescence microscopy (A. Taylor, unpubl.).

Bacterial production was measured on the P0704 cruise by the leucine incorporation technique (Kirchman et al., 1985; Simon and Azam, 1989). Seawater samples were inoculated with 20 nM of [4,5-^{3}H]-L-leucine and incubated at *in situ* temperature for 1 h. The incubations were terminated via addition of trichloroacetic acid (TCA; 5% final concentration). Control samples were prepared by adding TCA to each vial prior to the addition of seawater. TCA-killed samples were placed in the refrigerator for at least 1 h prior to filtration on 0.2 μm pore size polycarbonate filters. The filters and the filtration

manifold were washed 2 times with 5% TCA. The filter was then removed, placed into a glass vial, and allowed to dry. 5 mL of Ecolite scintillation cocktail was added to each vial, shaken, and placed at room temperature for at least 1 h before analysis on a Beckman Coulter LS6000TA liquid scintillation counter.

Vertical carbon fluxes for the 2006 cruise were determined from the ^{234}Th–^{238}U disequilibrium method (Pike et al., 2005) integrated over the upper 100 m and from the C:^{234}Th ratios on particles collected by an *in situ* pump at 100-m depth (Stukel et al., 2011). For P0605 Cycle 1, we used a steady state physical model that incorporated an upwelling rate of 1 m d^{-1}, but allowed upwelling rates to range from negligible to 2 m d^{-1} to assess measurement uncertainty. For the 2007 cruise, POC flux rates were determined from ^{234}Th–^{238}U disequilibrium and the average C:^{234}Th on particles collected with both sediment traps and *in situ* pumps at 100-m depth.

2.3. Model solutions

In addition to measured rates and mass-balance equalities, the model solutions were further limited by a set of 38 biological constraints including bounds on respiration, gross growth efficiency and excretion for each group, as well as assimilation, maximum ingestion and maximum clearance rates for the grazers (see Appendix 1). Except for the measured net rates of change of phytoplankton, we assumed that net accumulation rates of other biological components varied within ± 10% of their average biomass d^{-1}.

Since, the typical L_2MN method (Vézina and Platt, 1988) tends to zero out particular values and does not generate any confidence intervals, we instead adopted an ensemble L_2MN approach similar to the L_2MN Monte Carlo method used by Stukel and Landry (2010). For each model input (measured rates and biomasses and steady-state equalities), we generated a normal distribution from the experimentally determined means (zero for unmeasured steady-state equalities) and variances. Inputs were then chosen randomly from these distributions to allow for simultaneous variation of each input. Once the input parameters had been randomly selected, the L_2MN solution was calculated using the algorithm of Vézina and Platt (1988) as implemented in

MATLAB by Jackson et al. (2001). When the system of equations is solvable (i.e. there are no inconsistencies amongst the equalities and inequalities) a full rank solution can be used (k = size of b), and the residual error will be zero. Often inconsistencies existed in the data (e.g. when the difference of growth and grazing was less than the rate of change of the phytoplankton population). In such cases we followed the method of Olsen et al. (2006) and chose the highest rank solution that could generate an answer that satisfied the inequality constraints. For each cycle and numerical experiment, we solved the L_2MN for at least 10,000 independent (randomly selected) sets of input parameters and determined our L_2MN solution as the median of these 10,000 sets, while also calculating 95% confidence intervals for all ecosystem flows. When solutions did not converge to within 5% of a stable value, we computed additional solution sets until convergence was reached.

To calculate a maximum likelihood solution that averages over the entire possible solution space, we used the MCMC method of Kones et al. (2009) adapted for use in MATLAB (code has been made available on CCE LTER Datazoo website) from the R-Code of Van den Meersche et al. (2009). We used a long jump length of 2000 to ensure that each solution vector was independent of the vector preceding it. After a burn-in period of 20 iterations, an additional 100 iterations were carried out to characterize the solution space. For each cycle and numerical experiment we ran the MCMC method at least 100 different times (convergence to within 5% was tested) using randomly selected input parameters (as we did for the L_2MN method). This generated a total of 10,000 solution vectors that account for both measurement error and under-determinacy error. We calculated the median of these 10,000 solution vectors as our MCMC solution and also used them to generate confidence intervals for all model outputs. When inconsistent parameter sets were drawn, we decreased the rank of the solution, thus allowing for an inexact solution to the system of equations, and followed the MCMC method of Van den Meersche et al. (2009) with the probability of accepting new solution sets based on the ratio of $p(\mathbf{q_1}):p(\mathbf{q_2})$, where:

$$p(\mathbf{q}) \propto e^{-\frac{1}{2}\sigma^{-2}(\mathbf{A'q}-\mathbf{b'})^T W^2 (\mathbf{A'q}-\mathbf{b'})}$$

To test the ability of the L_2MN and MCMC methods to predict ecosystem flows accurately, we conducted a series of numerical experiments in which we withheld a particular *in situ* rate measurement (^{14}C-PP, microzooplankton grazing, mesozooplankton grazing, vertical POC flux, or bacterial production) from the dataset and solved the inverse problem with both the L_2MN and MCMC as described above. By varying the input parameters based on the assumption of a normal distribution, we generated medians and 95% confidence intervals that could then be compared to the experimental field measurements. Since our model remains more constrained than most inverse ecosystem models with one measurement input withheld, we also conducted experiments with pairs of measurements withheld and with only primary production specified as a measured rate.

2.4. Error sources

Inverse ecosystem modeling uncertainty can arise from three distinct sources: measurement error, solution error (underdetermined system), and structural error. Structural error is difficult to assess due to the vast array of possible model constructions, but we could address the other two errors in our analysis. To assess the role of measurement error, we ran MCMC simulations for 100 different random sets of measurement inputs to each cycle (as above) and 1000 iterations for each set of random measurements. We then computed the mean for each set of random measurements, and used the variability between these mean solutions to compute a coefficient of variation (CV) due only to measurement uncertainty for each flow. We then

determined the solution (underdeterminacy) error by using the MCMC to sample the solution space for the exact measurement values of each cycle, and used the resultant set of solutions to compute a CV for each flow derived only from model underdeterminacy.

2.5. Statistical analyses

We used two separate statistical analyses to assess the efficacy of each solution method in retrieving withheld measurements from our numerical experiments. First we computed the sum of the root mean squared error (RMSE) to assess the deviation of the model predicted rates from the withheld measurements (after log-transformation to normalize the deviations) for each numerical experiment. While this approach assesses the ability of the median to recover the withheld measurement, it does not take into account the width of the error bounds, which reflects how well the variable has been estimated. To assess the degree to which each method accurately depicts the true uncertainty in its results, we also calculated double the fraction of the modeled distribution that was less than the measured value (if the median modeled value was greater than the measured value). This corresponds to the p-value at which, given the modeled distribution of the predicted measurement, our confidence interval would fail to bracket the actual *in situ* measurement. For instance, a p-value of 0.02 (corresponding to 1% of the model distribution falling below the measured value) would suggest that the 98% confidence intervals of the modeled output would exclude the true measurement value. We then computed the geometric mean of this p-value for each numerical experiment (to determine average statistics encompassing all cycles). For the few instances where the measured value fell completely outside the simulated model distributions, we used the means and standard deviations from the model to construct a normal distribution that was used to estimate the p-value. Since distributions of RMSE and p-statistic were not normal, we used bootstrapping (sampling with replacement) techniques to test significance at the 95% confidence level.

3. Results

3.1. Predictions of withheld data

Results of the numerical experiments to assess the predictive capabilities of the L_2MN and MCMC solution schemes are shown in Fig. 2 and Table 2. In the full dataset, the measured rate of change in phytoplankton concentration (ΔP) is equal to the difference between net production (^{14}C-PP) and the sum of protozoan grazing, mesozooplankton grazing, and phytoplankton loss to detritus (PHYtoDET). Thus, the one unconstrained flow, PHYtoDET, is fixed by the measured rates. When the ^{14}C-PP measurement is withheld from model input, the difference between phytoplankton net growth and loss to detritus is set by the phytoplankton rate of change and grazing measurements. The individual magnitudes of phytoplankton net growth and loss to detritus, however, are poorly constrained since we have allowed phytoplankton to grow at rates of up to two doublings per day with no explicit constraint on non-grazing mortality. To minimize total system flow, the L_2MN consistently sets a low PHYtoDET rate of essentially zero, thus making ^{14}C-PP roughly equal to the sum of ΔP and grazing (Fig. 2a), which fits the measurement data well. In contrast, by sampling all possible solutions that satisfy the system constraints (or lack of them), the MCMC method leads to high flux estimates for both primary production and PHYtoDET (Fig. 2b). These two fluxes are highly correlated but poorly predicted when the ^{14}C-PP measurement is withheld.

The grazing rates of protozoans and mesozooplankton on phytoplankton are likewise constrained by the balance of ΔP, PHYtoDET and ^{14}C-PP. When the measured ^{14}C-PP rates are provided, both solution schemes predict the grazing terms reasonably well, with the *in*
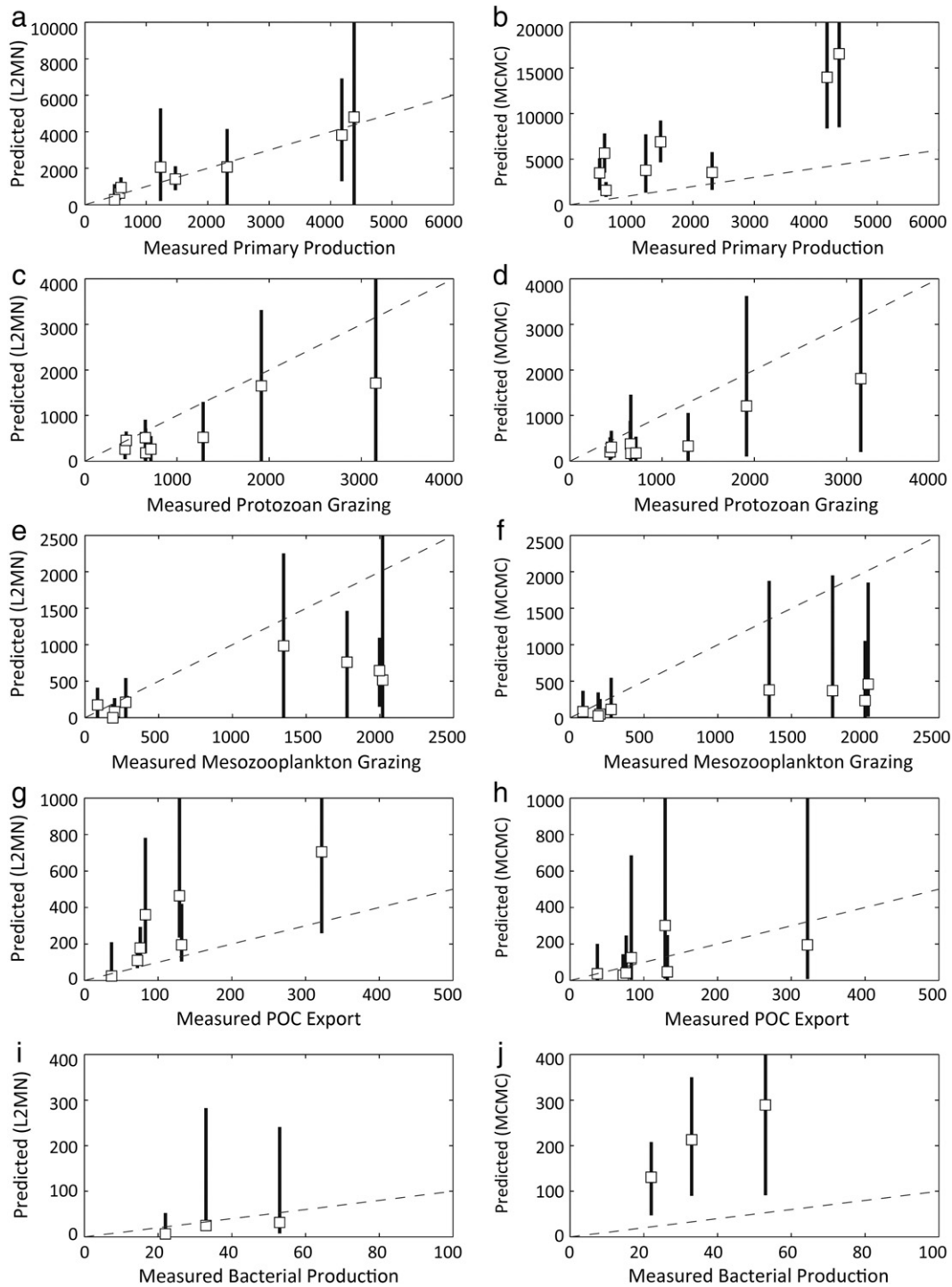
Fig. 2. Model estimation of measured parameters. Each measured rate was withheld from a model implementation in order to assess the model's ability to predict the measurement. Panels a, c, e, g, and i show the L_2MN predictions. Panels b, d, f, h, and j show the MCMC predictions. Panels a and b are the predictions of ^{14}C-PP when ^{14}C-PP is not used as a model input. Other measurements are protozoan grazing (c,d), mesozooplankton grazing (e,f), POC export (g,h) and bacterial production (i,j). *In situ* measurements are on the x-axis and model predictions are on the y-axis. All units are mg C m^{-2} d^{-1}. Dotted line is the 1:1 line. Error bars are 95% confidence limits.

situ measurements typically falling within the 95% confidence intervals for both the L_2MN and the MCMC. Nonetheless, both solution schemes underestimate grazing rates (Fig. 2c–f), especially when they are large.

Vertical POC flux and bacterial production are comparatively less constrained by the data. The L_2MN method consistently overestimates POC export, reflecting its tendency to shunt energy out of

the system as rapidly as possible (Fig. 2g). For three of the seven cycles (POC flux was not measured on P0605, Cycle 3), measured export values fall below the 95% confidence intervals of the L_2MN, and for six of the seven cycles the 67% confidence intervals fail to bracket the measurements. In comparison, the MCMC method predicts carbon export relatively well, with 95% confidence intervals bracketing the actual measurement values for all seven cycles and

Table 2

Summary statistics for numerical experiments in which measurement inputs were withheld. First column shows the measurement being predicted. These were always withheld from the model. PP = primary production, MicroGr = protozoan grazing, MesoGr = mesozooplankton grazing, BP = bacterial production. Mean is the means for all types of numerical experiments (one measurement withheld, two withheld, or PP as the only input). Second column shows other measurements that were withheld. Dash implies that the measurement being predicted was the only one withheld. 'All-PP' implies that primary production and mass balance constraints are the only model equalities. Third and fourth columns are the root mean squared error of the L_2MN and MCMC methods, respectively (after log-transformation to normalize the deviations). In two numerical experiments the median predicted for MesoGr by the L_2MN method was 0, leading to an infinite RMSE after log-transformation. Fifth and sixth columns are the geometric mean of the p-values (see Methods). p-values can vary from near 0 to 1, with a p-value of 1 implying that the model exactly predicts the measurement. A p-value of 0.05 would suggest that the actual measurement value lies on the 95% confidence level of the model's predicted distribution. Shown in bold are values for which the L_2MN solution was significantly different from the MCMC at the 95% confidence level.

Measurement	Withheld	RMSE		p-stat	
		L_2MN	MCMC	L_2MN	MCMC
PP	–	**0.14**	**0.64**	**0.582**	**1.2E−03**
MicroGr	–	**0.31**	**0.41**	0.097	0.072
MicroGr	MesoGr	0.19	0.28	0.077	0.140
MicroGr	POCFlux	0.38	0.43	0.056	0.081
MicroGr	All-PP	0.19	0.26	0.067	0.173
MesoGr	–	Inf	0.66	0.062	0.077
MesoGr	MicroGr	**0.30**	**0.55**	7.7E−05	0.103
MesoGr	POCFlux	Inf	0.59	0.070	0.082
MesoGr	All-PP	**0.30**	**0.57**	3.3E−05	0.075
POCFlux	–	0.39	0.30	**0.051**	**0.537**
POCFlux	MicroGr	**0.43**	**0.33**	**0.038**	**0.490**
POCFlux	MesoGr	0.48	0.28	**0.014**	**0.578**
POCFlux	All-PP	**0.48**	**0.31**	**0.028**	**0.515**
BP	–	**0.34**	**0.77**	**0.595**	**6.6E−04**
BP	All-PP	**0.52**	**0.74**	**0.223**	**0.010**
Mean	Single	Inf	0.56	**0.137**	**0.028**
Mean	Single[a]	Inf	0.53	0.088	0.073
Mean	Double	Inf	0.43	**0.015**	**0.173**
Mean	All-PP	0.36	0.46	**0.006**	**0.129**

[a] The second mean for only one measurement withheld is the mean excluding numerical experiments withholding ^{14}C-PP.

with even 50% confidence intervals bracketing the actual measurement value for five of the seven cycles (Fig. 2h). Bacterial production (BP), measured only on the P0704 cruise, was found to be surprisingly low, with BP:[14]C-PP ratios of 0.04, 0.04, and 0.01 for Cycles 1 through 3, respectively. By minimizing DOC production, the L_2MN solution predicted the low BP rates with excellent accuracy (Fig. 2i). The MCMC consistently overestimated BP by about a factor of five (Fig. 2j).

Since even with one measurement withheld, our ratio of constraints to flows (variables) remains higher than in most other inverse ecosystem models, we also conducted numerical experiments in which we withheld two measurements at a time and numerical experiments in which primary production was the only measurement input. The models continued to perform similarly in their predictions of grazing, while the L_2MN was a better predictor of bacterial production and the MCMC was a better predictor of carbon export. However, increasing underdeterminacy led to a stark difference in the uncertainty estimates of the two methods. While the error bounds suggested by the L_2MN method remained the same or even shrank as additional measurements were withheld, the MCMC error limits increased.

To test the significance of these trends we computed two statistics: The RMSE compares how closely the median model value predicts the withheld measurement, while the p-statistic compares how accurately the modeled probability distribution depicts our actual knowledge of the withheld measurement (including both the median value and uncertainty). With respect to RMSE, the L_2MN

typically performed better than the MCMC method, suggesting that its predictions were usually closer to the true measurement value. However, when averaged across all measurements, the difference between the RMSE of the L_2MN and MCMC methods was not statistically significant. For mesozooplankton grazing, the L_2MN method predicted a median value of 0 mg C m^{-2} for P0704-2, which led to an infinite RMSE after log-transformation. When this cycle is removed, the L_2MN outperformed the MCMC at predicting mesozooplankton grazing (RMSE = 0.38 and 0.61, respectively), when compared across all numerical experiments with a single measurement withheld (RMSE = 0.32 and 0.54), and with all numerical experiments except for predicting primary production (RMSE = 0.36 and 0.51). However, the difference was not statistically significant for mesozooplankton grazing alone, or when everything except primary production was averaged. By contrast, the MCMC method performed better than the L_2MN method with respect to the p-statistic, except when the large errors between MCMC predicted PP and ^{14}C-PP were included. In particular, the MCMC continued to perform relatively well even as model underdeterminacy increased, while the L_2MN performance dropped significantly. When PP was the only model input, the MCMC predictions included the measurements at a p-value of 0.129, while the L_2MN predictions would have excluded the measurements at a p-value of only 0.006.

3.2. Composite ecosystem indices

To further compare the L_2MN and MCMC methods we solved the LIM with all measurement inputs (Tables 3 and 4) and computed composite ecosystem indices for the different ecosystem states encountered on our cruises. While these indices are difficult to measure in the field, they can be easily calculated from ecosystem models and offer insights into the biases of each solution method.

One composite index is the gross growth efficiency (GGE) for all heterotrophic compartments (Fig. 3). The MCMC method gave GGE estimates of 17–24% for all grazer groups. The L_2MN method, however, predicted low GGEs of 10–17% for both protozoan groups, but a high GGE of >38% for mesozooplankton. The bacterial GGEs predicted by the two methods were surprising. The MCMC method predicted high GGEs (22–26%) when bacterial production was not measured on the P0605 cruise and low GGEs (5–7%) when it was measured (Fig. 3d). The GGE solutions from the L_2MN showed the opposite pattern. The MCMC solutions for bacterial GGEs on the P0704 cruise were driven to low values by the extremely low bacterial production measurements. The L_2MN predicted low GGEs when bacterial production was not measured, thus allowing bacteria to respire most of the carbon they consumed (shunting it out of the ecosystem to minimize the L_2 norm). Conversely, high GGE was predicted when bacterial production was measured because of consistently low DOC production in the L_2MN solutions.

To compare the overall utilization of primary production by different ecosystem components, we computed the amount of energy in three alternate pathways: the classical food chain (defined as the direct flow of phytoplankton, both living and detrital, to mesozooplankton), the multivorous food chain (defined as the flow reaching the mesozooplankton that stems from protozoans grazing on phytoplankton), and the microbial loop (defined as the sum of bacterial respiration and the portion of protozoan respiration supported by BP). The classical food chain was particularly well constrained. Although variability of the pathway was high (ranging from <20% to >140% of ^{14}C-PP), the L_2MN and MCMC strongly agreed on the magnitude for each set of ecosystem experiments (Fig. 4a). The solution methods differed strongly, however, in representing the multivorous food chain and the microbial loop, with the L_2MN consistently predicting less carbon utilization by both pathways. The L_2MN predicted low carbon flow from protozoans to mesozooplankton (Fig. 4b) due to

Table 3
MCMC solutions to the inverse ecosystem model. GPP = gross primary production. Fluxes are shown as SOURCEtoSINK, in units of mg C m^{-2} d^{-1}. Compartments are phytoplankton (PHY), heterotrophic nanoflagellates (HNF), microzoo-plankton (MIC), mesozooplankton (MES), bacteria (BAC), detritus (DET), and dissolved organic carbon (DOC). RES denotes respiration of the living compartments, while MEStoEXT and DETtoEXT are flows from mesozooplankton to higher trophic levels and from detritus that sink out of the euphotic zone, respectively. Median and 95% C.L. Column headings indicate cruise (e.g., 0605) and cycle number (e.g., 1).

	0605 Cycle1	0605 Cycle2	0605 Cycle3	0605 Cycle4	0605 Cycle5	0704 Cycle1	0704 Cycle2	0704 Cycle3
GPP	9100 (4980–14,061)	1284 (899–1721)	10,458 (6815–17,160)	3125 (2178–4215)	1017 (600–1623)	2995 (1014–5929)	1344 (899–2170)	5005 (2890–8031)
PHYtoRES	2939 (778–6391)	426 (117–812)	3545 (960–7763)	1032 (286–1968)	342 (91–741)	1137 (226–2860)	529 (146–1025)	2107 (728–4133)
PHYtoHNF	898 (46–2400)	221 (17–516)	1232 (36–3370)	364 (35–624)	174 (10–332)	170 (6–512)	280 (39–427)	185 (13–284)
PHYtoMIC	623 (27–2165)	147 (7–450)	979 (26–4917)	252 (15–678)	247 (16–760)	162 (6–619)	350 (65–603)	802 (197–1054)
PHYtoMES	1089 (51–2274)	139 (32–275)	1326 (28–2816)	252 (96–353)	70 (6–169)	1735 (600–3271)	98 (5–211)	1365 (298–2652)
PHYtoDET	1221 (78–3059)	134 (3–391)	1638 (77–5185)	135 (3–650)	233 (10–590)	145 (3–1423)	8 (0–349)	710 (11–2635)
PHYtoDOC	1643 (304–2895)	247 (54–360)	1775 (265–3308)	619 (151–884)	178 (31–324)	230 (37–843)	126 (18–375)	171 (50–651)
HNFtoMIC	330 (17–1135)	63 (3–224)	443 (26–1344)	85 (4–300)	56 (5–101)	141 (7–412)	52 (3–115)	78 (4–152)
HNFtoMES	255 (11–965)	52 (2–196)	238 (8–954)	74 (3–274)	29 (2–53)	112 (5–496)	63 (4–154)	227 (18–700)
HNFtoRES	749 (275–1501)	137 (59–267)	915 (373–1819)	195 (84–393)	124 (53–266)	319 (106–779)	137 (62–280)	436 (129–1183)
HNFtoDET	725 (182–1796)	148 (39–400)	775 (233–1690)	210 (55–577)	129 (33–352)	299 (64–984)	140 (39–375)	352 (71–1113)
HNFtoDOC	415 (143–904)	78 (31–167)	494 (192–1016)	109 (44–238)	69 (27–165)	161 (51–407)	71 (31–151)	196 (55–469)
MICtoRES	805 (303–1793)	132 (56–261)	1204 (433–3313)	211 (95–417)	158 (67–371)	286 (99–761)	166 (72–335)	523 (206–1264)
MICtoMES	602 (176–1495)	106 (35–257)	770 (249–2057)	163 (59–391)	113 (36–178)	213 (67–625)	136 (50–308)	353 (137–854)
MICtoDET	850 (210–2463)	147 (37–387)	1162 (284–3694)	251 (67–615)	183 (43–516)	264 (60–922)	186 (48–453)	431 (122–1151)
MICtoDOC	448 (157–1098)	75 (29–166)	655 (222–1872)	119 (49–254)	90 (34–225)	142 (47–392)	85 (36–177)	225 (96–490)
MEStoRES	790 (326–1541)	128 (65–243)	947 (429–1771)	197 (99–379)	116 (53–235)	667 (268–1656)	135 (64–278)	801 (339–1744)
MEStoDET	862 (228–1697)	145 (41–336)	826 (279–1589)	224 (64–508)	111 (31–284)	920 (267–2042)	140 (39–351)	764 (226–1544)
MEStoDOC	432 (167–904)	71 (32–148)	505 (218–1000)	107 (48–225)	63 (26–143)	328 (130–733)	69 (30–149)	336 (156–625)
BACtoRES	2626 (1231–4433)	407 (234–600)	3332 (1732–6801)	794 (419–1146)	374 (205–646)	958 (421–2058)	392 (213–720)	1108 (499–1707)
BACtoHNF	106 (6–204)	55 (3–116)	98 (5–190)	48 (3–87)	45 (2–100)	25 (1–66)	18 (1–71)	33 (1–74)
BACtoMIC	659 (103–1406)	71 (5–180)	820 (150–2058)	198 (30–312)	59 (3–173)	56 (1–231)	17 (1–87)	72 (2–186)
DETtoHNF	1536 (127–3626)	210 (12–655)	1522 (100–3507)	295 (17–1006)	187 (11–585)	863 (138–2305)	174 (9–558)	1137 (195–2608)
DETtoMIC	914 (41–3619)	139 (6–532)	1215 (53–4926)	177 (8–768)	151 (6–584)	479 (26–1871)	138 (6–593)	647 (30–1915)
DETtoMES	476 (19–1884)	93 (4–407)	366 (10–1744)	127 (5–585)	154 (8–454)	214 (8–1150)	114 (5–442)	363 (14–1634)
DETtoDOC	268 (10–1456)	42 (2–201)	514 (17–3219)	54 (2–264)	56 (2–275)	66 (2–392)	30 (1–154)	69 (2–411)
DOCtoBAC	3393 (1589–5618)	535 (309–777)	4231 (2217–8144)	1044 (543–1424)	485 (270–819)	1026 (469–2190)	421 (234–780)	1175 (531–1804)
DETtoEXT	256 (12–497)	56 (9–111)	388 (13–3028)	122 (61–178)	73 (4–145)	27 (0–99)	14 (0–42)	27 (0–180)
MEStoEXT	470 (181–1096)	81 (20–181)	581 (229–1396)	114 (45–262)	70 (9–180)	400 (132–1073)	90 (18–242)	499 (194–1160)

Table 4

L_2MN solutions to the inverse ecosystem model. Abbreviations as in Table 3. Median and 95% C.L.

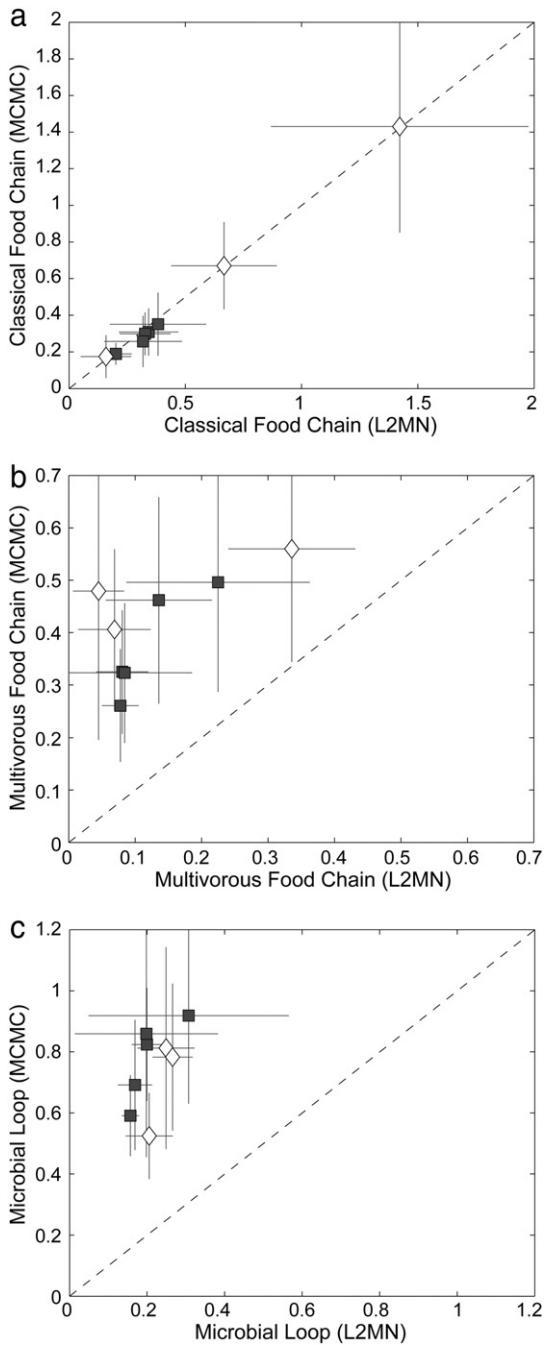|	0605 Cycle1	0605 Cycle2	0605 Cycle3	0605 Cycle4	0605 Cycle5	0704 Cycle1	0704 Cycle2	0704 Cycle3
GPP	5986 (3475–8440)	817 (758–981)	6613 (5659–11,276)	2099 (1742–2470)	697 (456–1157)	2051 (591–3689)	948 (748–1185)	3531 (2099–5213)
PHYtoRES	1502 (862–2160)	212 (189–262)	1691 (1409–2943)	508 (417–605)	187 (121–273)	534 (154–892)	247 (196–303)	911 (542–1351)
PHYtoHNF	911 (388–1615)	205 (101–333)	1382 (321–3880)	315 (192–470)	222 (67–347)	226 (50–450)	312 (204–433)	284 (259–293)
PHYtoMIC	904 (381–1612)	204 (98–333)	1327 (150–3732)	314 (190–469)	222 (67–936)	200 (32–426)	312 (204–449)	792 (238–1066)
PHYtoMES	1228 (382–2325)	169 (60–293)	1542 (0–3440)	258 (103–350)	74 (0–191)	1753 (589–3265)	93 (0–199)	1466 (464–2677)
PHYtoDET	330 (0–2608)	0 (0–291)	0 (0–4553)	60 (0–652)	207 (0–656)	0 (0–574)	0 (0–112)	0 (0–1892)
PHYtoDOC	209 (68–356)	22 (11–37)	198 (84–455)	94 (41–124)	10 (7–112)	31 (10–327)	19 (11–74)	73 (35–436)
HNFtoMIC	18 (0–105)	0 (0–17)	13 (0–459)	0 (0–5)	0 (0–4)	54 (0–101)	0 (0–0)	45 (0–119)
HNFtoMES	100 (0–216)	29 (0–65)	146 (0–694)	44 (17–76)	40 (10–55)	0 (0–0)	61 (33–91)	23 (0–146)
HNFtoRES	815 (290–1385)	147 (82–226)	1016 (366–2537)	244 (136–367)	180 (92–268)	284 (61–483)	212 (61–292)	412 (121–800)
HNFtoDET	121 (57–204)	24 (14–40)	186 (57–502)	41 (25–62)	29 (17–95)	45 (19–155)	35 (24–103)	61 (39–172)
HNFtoDOC	117 (49–198)	22 (13–35)	154 (55–414)	36 (22–53)	28 (17–41)	43 (20–70)	34 (22–46)	60 (37–117)
MICtoRES	820 (294–1389)	149 (83–227)	1005 (340–2423)	244 (137–367)	184 (104–362)	280 (61–495)	211 (61–295)	638 (205–1144)
MICtoMES	130 (13–250)	28 (12–64)	194 (36–891)	42 (23–76)	41 (15–179)	53 (0–100)	61 (33–95)	113 (0–212)
MICtoDET	122 (59–204)	24 (14–41)	185 (53–602)	41 (25–62)	29 (17–221)	45 (19–154)	35 (24–103)	95 (44–315)
MICtoDOC	118 (50–199)	22 (13–35)	153 (52–440)	36 (22–53)	29 (17–155)	42 (20–72)	34 (22–47)	94 (41–165)
MEStoRES	719 (259–1211)	112 (67–161)	860 (297–1505)	169 (89–266)	114 (53–174)	678 (257–1130)	127 (20–176)	696 (279–1161)
MEStoDET	180 (74–307)	27 (17–40)	217 (69–411)	41 (21–65)	28 (16–62)	197 (74–1070)	31 (19–43)	178 (84–540)
MEStoDOC	178 (67–301)	27 (16–38)	213 (69–366)	41 (21–65)	27 (15–44)	186 (79–392)	30 (8–42)	174 (82–288)
BACtoRES	626 (340–987)	92 (74–127)	732 (541–1928)	199 (155–256)	114 (66–203)	266 (129–463)	102 (72–143)	394 (225–702)
BACtoHNF	23 (0–64)	4 (0–14)	35 (0–194)	7 (0–26)	3 (0–20)	40 (0–69)	15 (0–33)	63 (0–79)
BACtoMIC	17 (0–56)	3 (0–14)	24 (0–299)	6 (0–26)	2 (0–10)	20 (0–150)	15 (0–33)	0 (0–109)
DETtoHNF	137 (0–807)	1 (0–84)	0 (0–812)	0 (0–185)	44 (0–183)	161 (20–330)	6 (0–33)	258 (64–804)
DETtoMIC	127 (0–803)	0 (0–82)	0 (0–705)	0 (0–185)	42 (0–178)	133 (11–305)	6 (0–32)	85 (0–571)
DETtoMES	218 (0–981)	30 (0–110)	0 (0–857)	53 (0–256)	101 (0–226)	0 (0–158)	78 (0–122)	27 (0–753)
DETtoDOC	0 (0–293)	0 (0–31)	0 (0–453)	0 (0–59)	18 (0–75)	0 (0–96)	3 (0–21)	0 (0–196)
DOCtoBAC	664 (370–1047)	98 (82–136)	774 (595–2089)	212 (171–272)	121 (71–251)	333 (180–610)	127 (93–169)	447 (268–769)
DETtoEXT	304 (83–527)	64 (15–114)	956 (359–3458)	126 (69–179)	77 (5–365)	48 (2–907)	20 (1–212)	76 (5–1147)
MEStoEXT	709 (248–1197)	104 (49–158)	849 (245–1506)	161 (82–259)	105 (30–170)	673 (236–1124)	119 (0–173)	688 (208–1158)

the previously mentioned low gross growth efficiencies of proto-zoans. Low activity of the microbial loop in the L_2MN solutions reflected the minimization of DOC production by phytoplankton and grazers (Fig. 4c).

The percentage direct contribution of phytoplankton to vertical POC flux (Fig. 5) compares the relative importance of sinking cells and grazing processes to carbon export for the two solution methods. While the MCMC method predicted relatively modest contributions
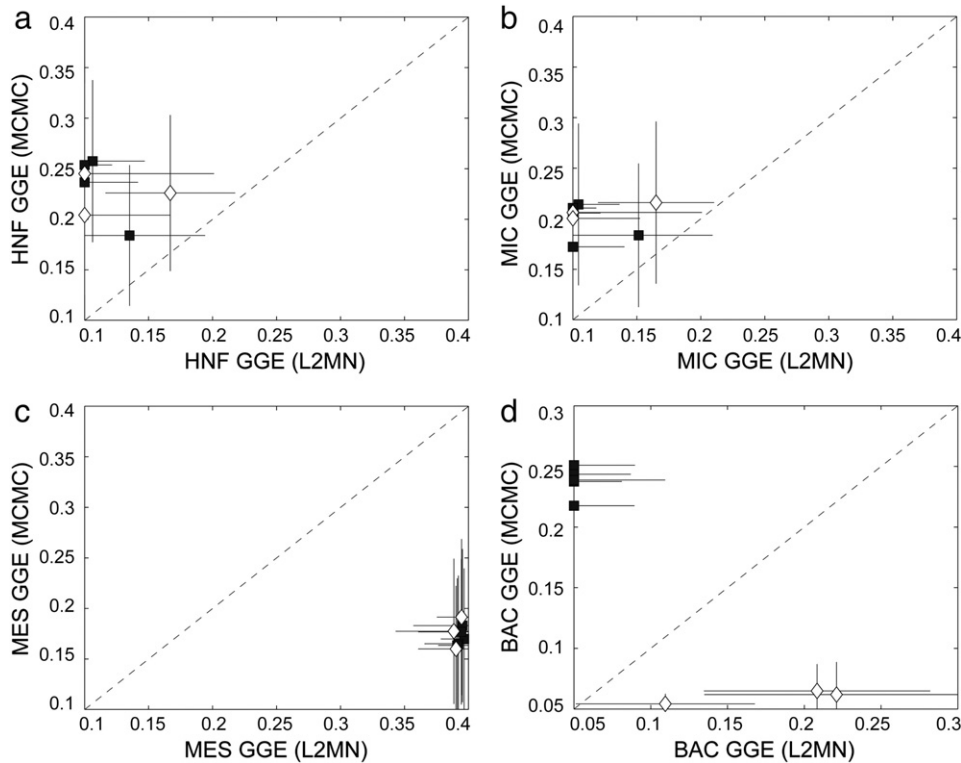
Fig. 3. Model predictions of gross growth efficiencies of heterotrophic nanoflagellates (a), microzooplankton (b), mesozooplankton (c), and bacteria (d). In panels, L_2MN predictions are on the x-axis and MCMC predictions are on the y-axis. Error bars are ± one standard deviation. Filled squares are from the P0605 cruise (May 2006), open diamonds are from the P0704 cruise (April 2007). Dotted line is the 1:1 line.
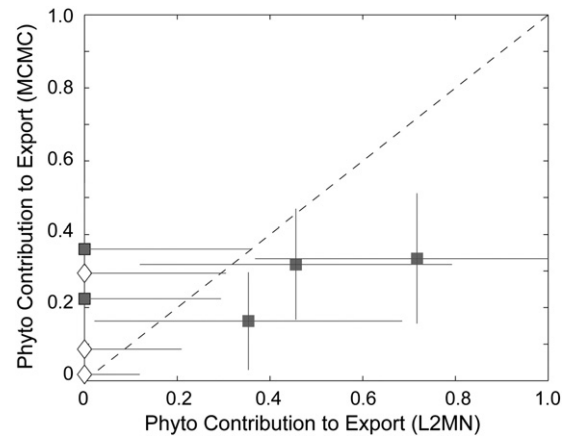
Fig. 5. Fraction of vertical carbon export derived from gravitational sinking of detrital phytoplankton, as determined by the L$_2$MN (x-axis) and the MCMC methods (y-axis). Filled squares are from the P0605 cruise, open diamonds are from the P0704 cruise. Dotted line is the 1:1 line.

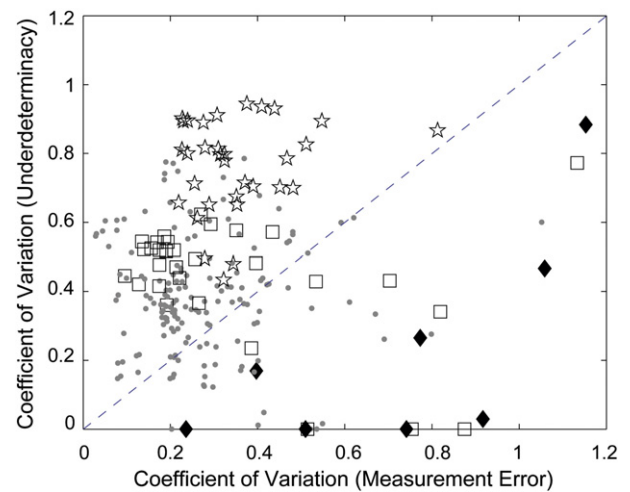**Fig. 6.** Comparison of the uncertainties due to underdeterminacy of the inverse model and to measurement errors. X-axis represents the coefficient of variation (CV) due to measurement error; y-axis is the CV due to underdeterminacy. Fluxes to and from detritus are highlighted. Flux from detritus (stars). Flux to detritus (squares). Black diamonds are detritus export (vertical POC flux). Dashed line is the 1:1 line.

Fig. 4. Comparisons of the relative strengths of the different food web components by the MCMC (y-axis) and L$_2$MN (x-axis) models. Panel a is the classical food chain defined as the sum of flows reaching the mesozooplankton that derive from phytoplankton without processing by protozoans; these include grazing on living and detrital phytoplankton. Panel b is the multivorous food chain, defined as the flow of carbon to mesozooplankton derived from protozoa. Panel c is the microbial loop, defined as the sum of the respiration of bacteria and the proportion of protozoan respiration that is derived from bacterial production. Flows for all cycles are normalized to ^{14}C-PP (mg C m^{-2} d^{-1}). Filled squares are from the P0605 cruise, open diamonds are from the P0704 cruise. Dotted line is the 1:1 line.

of gravitational phytoplankton sinking to export (1.7–36%), the L$_2$MN predicted widely varying contributions from 0% for five of the eight cycles to over 70% on P0605-5.

3.3. Error sources

For the majority (75%) of our ecosystem flows, solution errors of the inverse model exceeded the measurement errors (Fig. 6). The

exceptions were typically flows that were either directly measured or highly correlated with the field measurements (e.g. mesozooplankton grazing and POC flux were measured at sea, GPP was highly correlated with ^{14}C-PP). With the exception of POC export, flows involving detritus exhibited particularly high solution errors. This was especially the case for flows from detritus to grazers due to our lack of a priori knowledge or experimental measurement of detrital fluxes.

4. Discussion

4.1. System characteristics of the CCE during springtime

Based on floristic patterns, Venrick (2002) divided the CCE region into a temporally variable coastal domain dominated by large phytoplankton and a more stable offshore domain dominated by smaller cells. Three of our experimental cycles (P0605-2, P0605-5, and P0704-2) clustered with the offshore domain while the other five were characteristic of the upwelling-influenced coastal area. In addition to significantly lower productivity rates (Table 1), the offshore
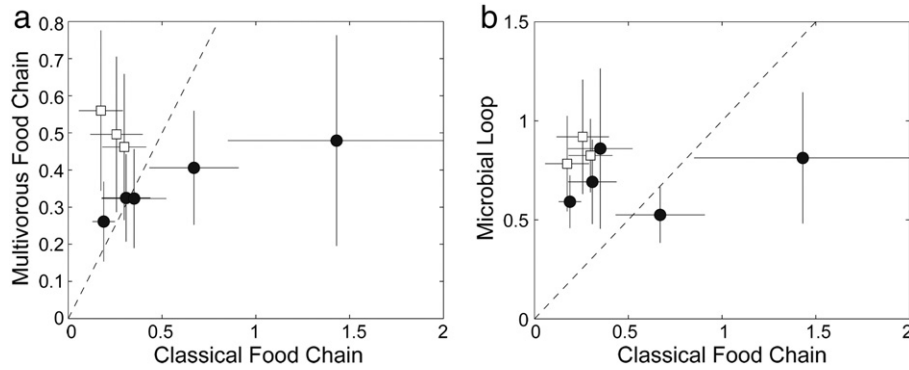
Fig. 7. Comparison of food webs. Panel a compares the multivorous food chain fluxes (y-axis) to the classical food chain fluxes (x-axis). Panel b compares the microbial loop (y-axis) to the classical food chain (x-axis). All food web fluxes are normalized to ^{14}C-PP (mg C m^{-2} d^{-1}). White squares are the offshore cycles. Black circles are the nearshore (upwelling influenced) cycles. Error bars are ±1 SD. Dashed line is the 1:1 line.

cycles were more similar to each other than the coastal cycles. All three exhibited weak functioning of the classical food chain, but relatively strong multivorous and microbial food webs (Fig. 7). In the more variable inshore domain, the strength of the classical food chain varied from 19 to 143% of ^{14}C-PP and was strongly and positively correlated with the multivorous food chain, suggesting that variations in the activity of the mesozooplankton community exert a major structuring influence on the coastal food web. The microbial loop also varied more in the nearshore area, ranging from 52 to 86% of ^{14}C-PP (compared to 78 to 92% for the offshore region), but its relatively high values for all experiments indicates that it is a major pathway of carbon flow throughout the region.

Given the apparent importance of mesozooplankton in structuring the coastal region, it is interesting to compare the modeled predictions of trophic level (TL) for this community component. Spatial patterns for the mesozooplankton TL were distinctly different on the two cruises (Fig. 8). On P0605, the predicted mesozooplankton TL was not strongly correlated with distance from shore. For P0704, however, the predicted increase in mesozooplankton TL with distance from shore agreed well with cruise measurements of feeding selectivity made on the copepod *C. pacificus* and the euphausiid *E. superba* (M. Décima, unpublished data). Across the range of ecosystem states sampled, the inverse model also predicted that direct gravitational sinking of phytoplankton contributed less to export (varying from 1.6 to 36% of total vertical flux in the MCMC model, Fig. 5), than grazing products, a

result that agrees well with pigment analyses from sediment traps deployed in the CCE region (M. Stukel, unpublished data).

4.2. Inferences from inverse ecosystem models

Since the pioneering work of Vézina and Platt (1988), inverse ecosystem modeling methods have been frequently used to compare different ecosystems (e.g. Daniels et al., 2006) or to infer unmeasured ecosystem flows (e.g. Jackson and Eldridge, 1992). However, investigators have not systematically utilized field data to assess the efficacy of inverse methods in approximating ecosystem structure. In the present study, we utilized rate measurements from two cruises of the CCE LTER program to test the accuracy of both the L$_2$MN and MCMC solution approaches for predicting cruise measurements and aggregate properties of the ecosystem. With both methods, we used a Monte Carlo technique of randomly varying the input parameters (measurements) for each cycle to generate an ensemble solution that incorporates measurement error. This ensemble solution is distinct from the typical exact solution that assumes perfect field measurements of ecosystem rates. In particular, it averages out some of the biases of the exact L$_2$MN solution while generating more realistic confidence intervals.

To evaluate the ability of inverse solution methods to predict known quantities, we sequentially withheld individual cruise measurements. Not surprisingly, both methods did a reasonable job of
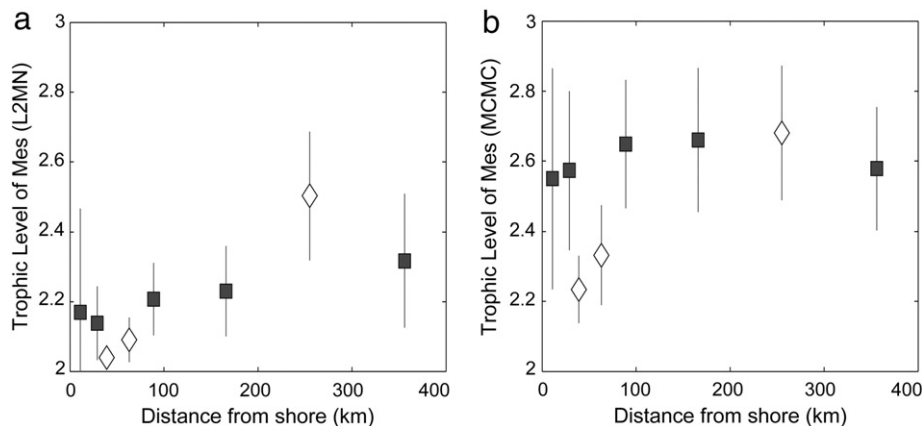
Fig. 8. Mesozooplankton trophic levels predicted by the MCMC (panel a) and the L$_2$MN (panel b) solution approaches. X-axis is the initial distance from shore. Black squares are P0605 and white diamonds are P0704. Error bars are ±1 SD.

estimating micro- and mesozooplankton grazing rates, as these flows are well constrained by related rate measurements (Fig. 2c–f). The slight underestimate of grazing rates by both methods arose because there were maximal constraints placed on grazing (see Appendix 1), while PHYtoDET losses were unconstrained. The L_2MN was far better than the MCMC method in predicting primary production (Fig. 2a, b). However, ^{14}C-PP is a relatively easy measurement to make, and we have not found a case for inverse models of pelagic ecosystems where it has not been a measured input. Furthermore, the MCMC's prediction of primary production might be significantly improved by placing a more stringent upper bound on phytoplankton growth rate. The maximum of two doublings per day that we allowed in the present analysis is unrealistically high for the relatively cold water of the CCE region during springtime.

Compared to production and grazing terms, model output of vertical POC flux and bacterial production are not well linked to the other measured rates since they depend on loosely constrained production and utilization of detritus and DOC, as well as other energy fluxes (like respiration) out of the ecosystem. Because they are the type of poorly constrained and highly coupled ecosystem flows that inverse models are often used to resolve, model predictions of vertical carbon flux and bacterial production are perhaps the more revealing aspect of our analyses. The MCMC method was much more accurate at predicting vertical carbon flux than the L_2MN approach (Fig. 2g,h). Confidence intervals of MCMC solutions consistently bracketed measurement values and solutions were typically less than a standard deviation away from the measured values. In contrast, the L_2MN solutions consistently overestimated POC export. Despite broad uncertainties for this flow (average CV for the 7 cycles was 1.98), 95% confidence intervals failed to bracket the measured values for three out of seven field experiments. The overestimate of POC export by the L_2MN may seem surprising given its tendency to maximize grazer respiration, but it follows logically from the method's minimization of flows through the system. Rather than

allow detritus to be reworked by grazers, the L_2MN exports most detritus directly out of the system.

The high L_2MN model uncertainty (typically of a similar magnitude to MCMC model uncertainty) suggested by the 95% confidence intervals in Fig. 2 is surprising, since the L_2MN method only incorporates measurement uncertainty, while the MCMC method includes uncertainty from both measurement inaccuracy and model underdeterminacy. This result is counterintuitive, since most model uncertainty was derived from underdeterminacy, and it arises because the coefficient of variation due only to measurement uncertainty was typically a factor of three larger in the L_2MN method than in the MCMC method. The high variability due to measurement uncertainty in the L_2MN model is due to the method's tendency to place flows at their upper or lower allowable bounds, thus leading to a strong correlation between a model flow and a single input parameter (Stukel and Landry, 2010). For instance, HNF GGE is typically exactly 10% and HNF respiration is maximized, hence HNF respiration is tightly correlated with measured protozoan grazing. By contrast, the MCMC method samples flows over their entire allowable ranges and hence depends on the multiple measurement inputs that affect its upper and lower bounds. If the uncertainty in different *in situ* measurements is uncorrelated, this will lead to decreased uncertainty in modeled flows.

In contrast to the export predictions, our few measurements of bacterial production by the leucine incorporation technique on the P0704 cruise were accurately predicted by L_2MN, but substantially overestimated by MCMC. This result needs to be interpreted with caution, however, because we assumed that utilizable (i.e. labile) DOC concentration remained constant in our experiments. Other studies have demonstrated significant net accumulation of DOC in waters advecting away from upwelling sources (Alvarez-Salgado et al., 2001; Wetz and Wheeler, 2004). During the P0704 studies in the upwelling system off of Point Conception, substantial net DOC accumulation may have occurred as a consequence of excess grazing over phytoplankton growth
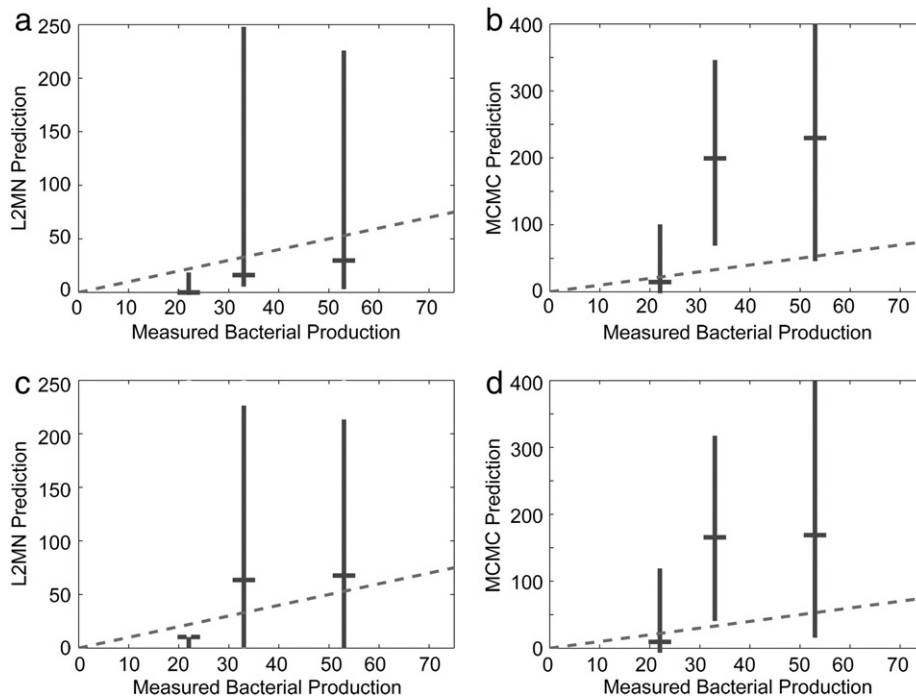

Fig. 9. Model predictions of bacterial production (mg C m^{-2} d^{-1}) with assumed DOC accumulation. Figures are comparable to Fig. 2i, j. Panels a and b assume labile DOC accumulation in the euphotic zone at a rate of 1% d^{-1}; panels c and d assume accumulation rates of 2% d^{-1}. Panels a and c are L_2MN predictions, while panels b and d are MCMC predictions.

(declining phytoplankton biomass; Landry et al., 2009), and disproportionately low rates of DOC utilization by bacteria (the 1–4% ratio of BP to ^{14}C-PP measured on P0704 is far below the typical marine average of 10–20%); (Ducklow, 2000). These low rates may have been due to low bacterial biomass in the spring season or to cold temperature depression of specific growth rates.

If we modify the model to allow for even slight net production of DOC (1 and 2% d^{-1}; Fig. 9), the L$_2$MN and MCMC do similarly well in predicting bacterial production, with the MCMC becoming increasingly better as DOC accumulation is increased, albeit with large uncertainty. Measured DO^{14}C production rates on springtime cruises in the CCE region as part of the CalCOFI (California Cooperative Oceanic Fisheries Investigations) Program average roughly 20% of ^{14}C-PP (R. Goericke, unpublished data) during 24-h incubations. For the L$_2$MN solution, our estimates of PHYtoDOC were low, ranging from 2.5 to 6.4% of ^{14}C-PP (mean 3.9%) across 8 experimental cycles. The MCMC solutions, by contrast, predicted phytoplankton DOC production rates of 7.4 to 43.8% (mean 31.3%), which compare more favorably with the measured DO^{14}C production rates on CalCOFI cruises.

Although respiration and GGE were not measured at sea, the distinct differences between the models' apportionment of grazer assimilation into alternate pathways of production, respiration and excretion allow for a useful comparison of the two solution methods. In a meta-analysis of laboratory experiments determining GGEs of flagellates, ciliates, dinoflagellates and crustaceans under food-replete conditions, Straile (1997) found little variability across taxa, with all having mean GGEs between 20 and 30%. The only size relationship was a slight (statistically insignificant) decrease in GGE with increasing size. This agrees well with mean GGE estimates of 23%, 20% and 17% for heterotrophic nanoflagellates, microzooplankton, and mesozooplankton, respectively, by the MCMC method. In contrast, the L$_2$MN predictions of substantially lower GGEs for protistan grazers (11% and 12% for nanoflagellates and microzooplankton, respectively) and much higher GGEs (39%) for the metazoans (Fig. 3) seem decidedly unrealistic given the expectation that the GGEs of larger consumers should be more sensitive to, and more likely to decrease with, food limiting conditions in ambient pelagic ecosystems compared to smaller protists.

4.3. L$_2$MN and MCMC solution biases

Biases of the L$_2$MN solution method are reasonably well understood from previous studies. In particular, L$_2$MN tends to zero out minor flows (Vézina et al., 2004) and minimize the largest flows in the model. It also often sets biological constraints (e.g. GGE) to either maximum or minimum allowed values and minimizes the production of DOC (Stukel and Landry, 2010). The MCMC is a much newer technique, and as such its inherent biases have not been as well characterized. By fully sampling the solution space, the MCMC avoids many of the problems associated with the L$_2$MN. In particular, the MCMC selects more likely values near the middle of allowable parameter ranges, rather than extremes (Fig. 3). However, we did identify one MCMC behavior that can affect flux interpretations in models with very slight structural differences. When sampling the solution space, the MCMC method typically apportions unconstrained flows equally between equivalent compartments. In our model construction, for example, carbon can flow from nanoflagellates to respiration, DOC, detritus, microzooplankton, or mesozooplankton. If we add another pathway to those alternatives, for instance, by dividing microzooplankton into two assemblages (larger flagellates and ciliates) that both feed on nanoflagellates, the total flow into the combined microzooplankton group will increase at the expense of the other outflows from nanoflagellates. This is simply due to the larger number of pathways that can be constructed. Thus, with the MCMC technique (as with the L$_2$MN), it may be critically important to use identical structures when comparing different ecosystems or states.

5. Conclusion

Using data from 8 field experiments across a range of conditions in the California Current Ecosystem, we found that inverse methods reasonably estimated *in situ* measurements when measurement errors are taken into account. While an ensemble version of the commonly used L$_2$MN technique clearly performed better than the newer MCMC method when primary production was not specified as a model input, the MCMC method more accurately depicted pelagic food web flows when primary production was provided. In particular, the MCMC method predicted vertical carbon flux, which depends on a number of unmeasured rates; hence the MCMC appears to do especially well in capturing the dynamics of the poorly resolved, but highly connected, detrital pool. This skill, combined with its ability to generate realistic confidence intervals even as the system becomes less constrained, suggests that it may perform well with the highly underdetermined ecosystems often encountered. Nevertheless, results from inverse modeling techniques are highly uncertain, as shown by coefficients of variation that often exceed 50%. Despite the fact that our simple model, with 10–12 equations and 29 flows, was far more constrained than most inverse models, most of the variability arose from the under-determinacy of the solutions, rather than from measurement uncertainties. We thus suggest that experimental programs focusing on food web fluxes expand the range of experimental measurements, in particular, measurements that assess the nature and fate of the detrital pool, which plays a large role in the model. Techniques that can ascertain the composition of sinking detritus (e.g. Lundsgaard and Olesen, 1997) and rates of particle feeding (e.g. Wilson et al., 2010) are especially necessary for increased understanding of POC flux in the pelagic ocean.

Appendix 1. Model equalities (Ax = b) and inequalities (Gx ≥ h)

Equalities

$$GPPtoPHY - PHYtoRES - PHYtoHNF - PHYtoMIC - PHYtoMES - PHYtoDET - PHYtoDOC = PhyAccRate$$

$$PHYtoHNF + BACtoHNF + DETtoHNF - HNFtoMIC - HNFtoMES - HNFtoRES - HNFtoDET - HNFtoDOC = 0 \pm 10\% \text{ Biomass}$$

$$PHYtoMIC + HNFtoMIC + BACtoMIC + DETtoMIC - MICtoRES - MICtoMES - MICtoDET - MICtoDOC = 0 \pm 10\% \text{ Biomass}$$

$$PHYtoMES + HNFtoMES + MICtoMES + DETtoMES - MEStoRES - MEStoDET - MEStoDOC - MEStoEXT = 0 \pm 10\% \text{ Biomass}$$

$$DOCtoBAC - BACtoRES - BACtoHNF - BACtoMIC = 0 \pm 10\% \text{ Biomass}$$

$$PHYtoDET + HNFtoDET + MICtoDET + MEStoDET - DETtoHNF - DETtoMIC - DETtoMES - DETtoDOC - DETtoEXT = 0$$

$$PHYtoDOC + HNFtoDOC + MICtoDOC + MEStoDOC + DETtoDOC - DOCtoBAC = 0$$

$$GPPtoPHY - PHYtoRES - PHYtoDOC = {}^{14}\text{C-PP}$$

$$PHYtoHNF + PHYtoMIC = protogr$$

$$PHYtoMES = mesogr$$

$$DETtoEXT = Export$$

$$DOCtoBAC - BACtoRES = BacProd$$

Inequalities:

Rate	Pop	Minimum	Maximum
GGE	HNF	10%[1]	40%[1]
	MIC	10%[1]	40%[1]
	MES*	10%[1*]	40%[1*]
	BAC	5%[2]	30%[2]
BasalResp	MES	$2.314 * 10^{(\log(9.1) + \log(2) * (T-25)/10)} * \text{Biomass}$ [3]	
Assim. Eff.	HNF	50%[4]	90%[4]
	MIC	50%[4]	90%[4]
	MES	50%[4]	90%[4]
Respiration	PHY	10%	55%[5]
	HNF	20% ingestion[6]	
	MIC	20% ingestion[6]	
	MES	20% ingestion[6]	
	BAC	20% ingestion[6]	
Excretion	PHY	2% NPP[6,7]	55% NPP[6,7]
	HNF	10% ingestion[8]	100% respiration[9]
	MIC	10% ingestion[8]	100% respiration[9]
	MES	10% ingestion[8]	100% respiration[9]
Max Ingestion	HNF		$24 * 10^{\log(0.6509) + \log(2.8) * (T-20)/10} * \text{Biomass}$ [10]
	MIC		$24 * 10^{\log(0.2375) + \log(2.8) * (T-20)/10} * \text{Biomass}$ [10]
	MES		$24 * 10^{\log(0.1816) + \log(2.8) * (T-20)/10} * \text{Biomass}$ [10]
Max Clearance Rate	HNF		$24 * 10^{\log(203,030) + \log(2.8) * (T-20)/10} * \text{Biovolume}$ [10]
	MIC		$24 * 10^{\log(176,830) + \log(2.8) * (T-20)/10} * \text{Biovolume}$ [10]
	MES		$24 * 10^{\log(112,780) + \log(2.8) * (T-20)/10} * \text{Biovolume}$ [10]
GPP	PHY	140%NPP[11]	260%NPP[11]
NPP	PHY		$3 * \text{PhyBiomass}$ [12]

[1] Straile (1997).

*Because metazooplankton have a basal respiration rate in addition to their activity-dependent metabolic cost, we have implemented a basal metabolic rate that must be paid first, and constrained their GGE such that only 10–40% of the remainder of their intake can contribute to production. Thus the true GGE of the metazooplankton in our model will range from slightly less than 10% to slightly less than 40%.

[2] Del Georgio and Cole (2000).

[3] Makarieva et al. (2008). An allometric and temperature specific basal metabolic relationship for copepods and krill was used to set a minimum metabolic rate for metazooplankton that is in addition to their ingestion rate specific constraints.

[4] Conover (1966).

[5] Falkowski et al. (1985).

[6] Vézina et al. (2000).

[7] Baines and Pace (1991).

[8] Vézina and Pace (1994).

[9] Vézina and Platt (1988).

[10] Hansen et al. (1997). To ensure that we used maximal clearance and ingestion rates and because grazers are often believed to aggregate at regions of high prey density, the maximum concentrations of prey and predator encountered on each cycle were utilized in setting maximum clearance and ingestion rates. Since epifluorescence microscopy often significantly underestimates ciliate concentrations and all dinoflagellates (both pigmented and aplastidic) are likely phagotrophic, we included autotrophic dinoflagellate biomass in the maximum microzooplankton biomass figures used for calculating clearance and ingestion rates. Biomass was converted to biovolume using conversion factors of 200 and 130 mg C cm^{-3} for heterotrophic nanoflagellates and microzooplankton, respectively (Menden-Deuer and Lessard, 2000) and a biomass to biovolume relationship determined for mesozooplankton in the CCE (Lavaniegos and Ohman, 2007).

[11] Bender et al. (1999), Laws et al. (2000), Hashimoto et al. (2005).

[12] Production rate based on a specific growth rate of 1.4 d^{-1} or two doublings per day (Calbet and Landry, 2004).

References

Alvarez-Salgado, X.A., et al., 2001. Off-shelf fluxes of labile materials by an upwelling filament in the NW Iberian Upwelling System. Progress in Oceanography 51 (2–4), 321–337.

Baines, S.B., Pace, M.L., 1991. The production of dissolved organic matter by phytoplankton and its importance to bacteria: patterns across marine and freshwater systems. Limnology and Oceanography 36 (6), 1078–1090.

Bender, M., Orchardo, J., Dickson, M.L., Barber, R., Lindley, S., 1999. In vitro O$_2$ fluxes compared with ^{14}C production and other rate terms during the JGOFS Equatorial Pacific experiment. Deep-Sea Research I 46 (4), 637–654.

Calbet, A., Landry, M.R., 2004. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. Limnology and Oceanography 49 (1), 51–57.

Conover, R.J., 1966. Assimilation of organic matter by zooplankton. Limnology and Oceanography 11 (3), 338–345.

Daniels, R.M., Richardson, T.L., Ducklow, H.W., 2006. Food web structure and biogeochemical processes during oceanic phytoplankton blooms: an inverse model analysis. Deep-Sea Research II 53 (5–7), 532–554.

del Georgio, P.A., Cole, J.J., 2000. Bacterial energetics and growth efficiency. In: Kirchman, D.L. (Ed.), Microbial Ecology of the Oceans. Wiley-Liss, Inc., pp. 289–325.

Ducklow, H., 2000. Bacterial production and biomass in the oceans. In: Kirchman, D.L. (Ed.), Microbial Ecology of the Oceans. Wiley-Liss, pp. 85–120.

Falkowski, P.G., Dubinsky, Z., Wyman, K., 1985. Growth-irradiance relationships in phytoplankton. Limnology and Oceanography 30 (2), 311–321.

Hansen, P.J., Bjornsen, P.K., Hansen, B.W., 1997. Zooplankton grazing and growth: scaling within the 2–2000-mu m body size range. Limnology and Oceanography 42 (4), 687–704.

Hashimoto, S., Horimoto, N., Yamaguchi, Y., Ishimaru, T., Saino, T., 2005. Relationship between net and gross primary production in the Sagami Bay, Japan. Limnology and Oceanography 50 (6), 1830–1835.

Jackson, G.A., Eldridge, P.M., 1992. Food web analysis of a planktonic system off Southern California. Progress in Oceanography 30 (1–4), 223–251.

Jackson, G.A., Niquil, N., Burd, A., Richardson, T.L., 2001. Ecosystem Modeling Group Software, Inverse model code. Texas A&M Univ.http://oceanz.tamu.edu/~ecomodel/Software/invmodel/invmodel.html.

Kirchman, D., Knees, E., Hodson, R., 1985. Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. Applied and Environmental Microbiology 49 (3), 599–607.

Kones, J.K., Soetaert, K., van Oevelen, D., Owino, J.O., Mavuti, K., 2006. Gaining insight into food webs reconstructed by the inverse method. Journal of Marine Systems 60 (1–2), 153–166.

Kones, J.K., Soetaert, K., van Oevelen, D., Owino, J.O., 2009. Are network indices robust indicators of food web functioning? A Monte Carlo approach. Ecological Modelling 220 (3), 370–382.

Landry, M.R., Ohman, M.D., Goericke, R., Stukel, M.R., Tsyrklevich, K., 2009. Lagrangian studies of phytoplankton growth and grazing relationships in a coastal upwelling ecosystem off Southern California. Progress in Oceanography 83, 208–216.

Lavaniegos, B.E., Ohman, M.D., 2007. Coherence of long-term variations of zooplankton in two sectors of the California Current System. Progress in Oceanography 75 (1), 42–69.

Laws, E.A., et al., 2000. Carbon cycling in primary production bottle incubations: Inferences from grazing experiments and photosynthetic studies using ^{14}C and ^{18}O in the Arabian Sea. Deep-Sea Research II 47 (7–8), 1339–1352.

Lundsgaard, C., Olesen, M., 1997. The origin of sedimenting detrital matter in a coastal system. Limnology and Oceanography 42 (5), 1001–1005.

Makarieva, A.M., et al., 2008. Mean mass-specific metabolic rates are strikingly similar across life's major domains: evidence for life's metabolic optimum. Proceedings of the National Academy of Sciences of the United States of America 105 (44), 16994–16999.

Menden-Deuer, S., Lessard, E.J., 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. Limnology and Oceanography 45 (3), 569–579.

Niquil, N., Jackson, G.A., Legendre, L., Delesalle, B., 1998. Inverse model analysis of the planktonic food web of Takapoto Atoll (French Polynesia). Marine Ecology Progress Series 165, 17–29.

Olsen, Y., et al., 2006. A comparative study of responses in planktonic food web structure and function in contrasting European coastal waters exposed to experimental nutrient addition. Limnology and Oceanography 51 (1), 488–503.

Pike, S.M., Buesseler, K.O., Andrews, J., Savoye, N., 2005. Quantification of ^{234}Th recovery in small volume sea water samples by inductively coupled plasma-mass spectrometry. Journal of Radioanalytical and Nuclear Chemistry 263 (2), 355–360.

Simon, M., Azam, F., 1989. Protein content and protein synthesis rates of planktonic marine bacteria. Marine Ecology Progress Series 51 (3), 201–213.

Soetaert, K., van Oevelen, D., 2009. Modeling food web interactions in benthic deep-sea ecosystems: a practical guide. Oceanography 22 (1), 128–143.

Steele, J.H., 2009. Assessment of some linear food web methods. Journal of Marine Systems 76 (1–2), 186–194.

Straile, D., 1997. 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 (6), 1375–1385.

Stukel, M.R., Landry, M.R., 2010. Contribution of picophytoplankton to carbon export in the equatorial Pacific: a re-assessment of food-web flux inferences from inverse models. Limnology and Oceanography 55 (6), 2669–2685.

Stukel, M.R., Landry, M.R., Benitez-Nelson, C.R., Goericke, R., 2011. Trophic cycling and carbon export relationships in the California Current Ecosystem. Limnology and Oceanography 56, 1866–1878.

Van den Meersche, K., Soetaert, K., Van Oevelen, D., 2009. xSample(): an R function for sampling linear inverse problems. Journal of Statistal Software, Code Snippets 30 (1), 1–15.

van Oevelen, D., et al., 2006. Carbon flows through a benthic food web: integrating biomass, isotope and tracer data. Journal of Marine Research 64 (3), 453–482.

van Oevelen, D., et al., 2010. Quantifying food web flows using linear inverse models. Ecosystems 13 (1), 32–45.

Venrick, E.L., 2002. Floral patterns in the California Current System off southern California: 1990–1996. Journal of Marine Research 60 (1), 171–189.

Vézina, A.F., Pace, M.L., 1994. An inverse model analysis of planktonic food webs in experimental lakes. Canadian Journal of Fisheries and Aquatic Sciences 51 (9), 2034–2044.

Vézina, A.F., Pahlow, M., 2003. Reconstruction of ecosystem flows using inverse methods: how well do they work? Journal of Marine Systems 40, 55–77.

Vézina, A.F., Platt, T., 1988. Food web dynamics in the ocean .1. Best estimates of flow networks using inverse methods. Marine Ecology Progress Series 42 (3), 269–287.

Vézina, A.F., et al., 2000. Export of biogenic carbon and structure and dynamics of the pelagic food web in the Gulf of St. Lawrence Part 2. Inverse analysis. Deep-Sea Research II 47 (3–4), 609–635.

Vézina, A.F., Berreville, F., Loza, S., 2004. Inverse reconstructions of ecosystem flows in investigating regime shifts: impact of the choice of objective function. Progress in Oceanography 60 (2–4), 321–341.

Wetz, M.S., Wheeler, P.A., 2004. Response of bacteria to simulated upwelling phytoplankton blooms. Marine Ecology Progress Series 272, 49–57.

Wilson, S.E., Steinberg, D.K., Chu, F.L.E., Bishop, J.K.B., 2010. Feeding ecology of mesopelagic zooplankton of the subtropical and subarctic North Pacific Ocean determined with fatty acid biomarkers. Deep-Sea Research I 57 (10), 1278–1294.