Journal of Marine Systems 91 (2012) 20–33

Contents lists available at SciVerse ScienceDirect

Journal of Marine Systems

journal homepage: www.elsevier.com/locate/jmarsys

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

Article history:
Received 5 April 2011
Received in revised form 1 September 2011
Accepted 9 September 2011
Available online 21 September 2011

Keywords:
Linear inverse model
Plankton
Trophic structure
Food web
Ecosystem dynamics
California Current System

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.

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,

---

[*] Corresponding author at: Horn Point Laboratory, University of Maryland, 2020 Horns Point Rd., Cambridge, MD, 21613, USA. Tel.: +1 815 258 3875.
E-mail address: mstukel@hpl.umces.edu (M.R. Stukel).

0924-7963/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmarsys.2011.09.004

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 tradeoffs 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 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.

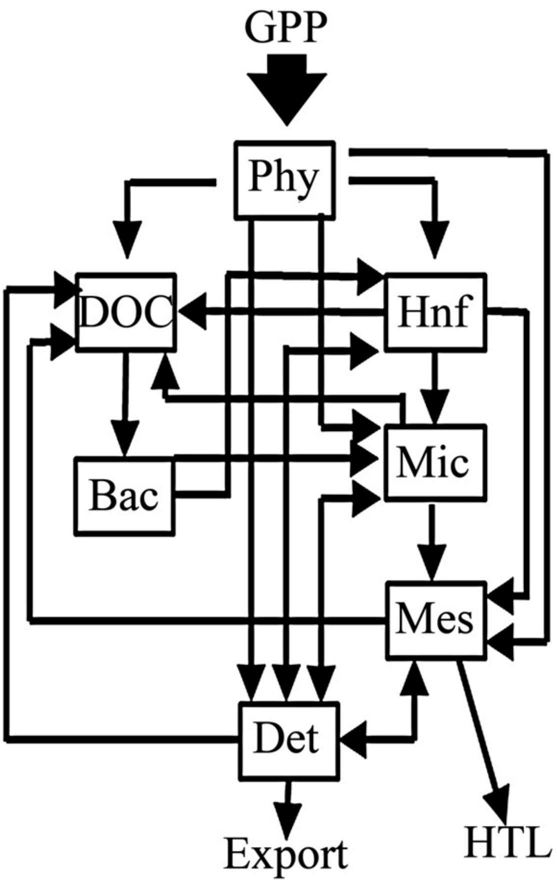

Fig. 1. Model structure. Model compartments are phytoplankton (Phy), heterotophic nanoflagellates (Hnf), microzooplankton (Mic), mesozooplankton (Mes), bacteria (Bac), detritus (Det), and dissolved organic carbon (DOC). GPP = gross primary production. Arrows show the direction of flow through the ecosystem. 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.

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