Progress in Oceanography 138 (2015) 546–558

Contents lists available at ScienceDirect

Progress in Oceanography

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

Impact of assimilating physical oceanographic data on modeled ecosystem dynamics in the California Current System

Kaustubha Raghukumar*, Christopher A. Edwards, Nicole L. Goebel, Gregoire Broquet, Milena Veneziani, Andrew M. Moore, Jon P. Zehr

Institute of Marine Sciences, University of California, Santa Cruz, CA 95064, USA

ARTICLE INFO

Article history:
Available online 4 February 2015

ABSTRACT

A method to model ocean ecosystems using data-constrained physical circulation estimates is investigated. Physical oceanographic data is assimilated into a Regional Ocean Modeling System implementation of the California Current System using an incremental 4-Dimensional Variational method. The resulting state estimate drives a complex, self-assembling ocean ecosystem model for the year 2003, and results are evaluated against SeaWiFS surface and CalCOFI subsurface observations and with ecosystem model output driven by an unconstrained physical model. While physical data assimilation improves correlation with observations, this method also drives elevated levels of phytoplankton standing stock, leading to a large bias particularly in regions of low mean concentration. We identify two causes for this increase: biological rectification of fluctuating vertical nutrient transport due to gravity wave generation at assimilation cycle initialization and increased nutrient variance on density surfaces. We investigate one and propose other possible remedies for these deleterious side-effects of this data assimilation method.

Introduction

One major development in oceanography over the last decade has been the practical realization of ocean observing systems that serve a multitude of uses such as monitoring upper ocean heat content at a global scale and improving coastal search and rescue operations in regional applications (Dexter and Summerhayes, 2010). Data collected from a wide variety of platforms, including satellites, drifters, moorings and HF RADAR sites, are made available to the public in near-realtime (see, for example, http://www.nodc.noaa.gov/access/). Such information also is assimilated routinely into oceanographic models to produce estimates of ocean variables that sensibly interpolate and extrapolate sparse data in space and time. Data assimilation methods for physical variables in the ocean are quite advanced today, and most ocean state estimates result from assimilation of physical data into ocean circulation models (Behringer and Xue, 2004). Accompanying these successes is an increasing interest in complementary biogeochemical information to support, for example, monitoring and management of hypoxia, harmful algal blooms, and marine fisheries (Jannasch et al., 2008).

Though less mature than its physical counterpart, data assimilation efforts for ocean ecosystem models also have been actively developed. Goals generally fall into two categories: estimation of otherwise poorly known biological model parameters (e.g., Matear, 1995; Lawson et al., 1995; Friedrichs, 2002; McGillicuddy et al., 1998a) and production of ocean ecosystem state estimates (e.g., Ishizaka, 1990; Natvik and Evensen, 2003). Reviews of biological data assimilation activities can be found in Gregg (2008) and Edwards et al. (2015). To date, these investigations have focused primarily on the assimilation of biological or chemical data into biogeochemical models. A more comprehensive approach involves the joint assimilation of physical and biological data into coupled physical and biological models (Song et al., 2012; Shulman et al., 2013). These methods are relatively complex and challenging, and their continuing development encourages consideration of alternate approaches to estimate ecosystem fields. One obvious and relatively simple strategy is to drive biogeochemical models with physical circulation and mixing estimates that have themselves been produced by physical data assimilation. This approach is attractive since physical ocean state estimates presently exist within many ocean observing systems and reanalyses, and it is reasonable at first consideration to expect that physical circulations, when already constrained by data, should result in improved ecosystem estimates.

A few of studies have tested this idea in the North Atlantic. Oschlies and Garçon (1999) assimilated remotely sensed sea

---

* Corresponding author at: Department of Oceanography, Naval Postgraduate School, 833 Dyer Road, Monterey, CA 93943, USA. Tel.: +1 831 656 7911, fax: +1 831 656 2712.

E-mail address: kraghuku@nps.edu (K. Raghukumar).

http://dx.doi.org/10.1016/j.pocean.2015.01.004
0079-6611/

surface height data into a coarse resolution model. While they found that greater mesoscale activity resulting from assimilation increased nutrient supply to the subtropical ocean, the resulting increase in primary production was not sufficient to match independent estimates. Anderson et al. (2000) carried out a series of experiments, considering both physical and biological assimilation in isolation and jointly. They stated that misalignments between physical and biological fields led to spurious biological production unless both physical and biological data were assimilated and adjusted to one another. Berline et al. (2007) showed that while surface chlorophyll estimates improved, vertical advective fluxes of nutrients showed a four- to sixfold increase in the mid-latitudes and sub-tropics, offering three potential explanations: (1) improved fidelity of the eddy-pumping process (McGillicuddy et al., 1998b); (2) misalignment in biological and physical fields (Anderson et al., 2000); and (3) adjustment of physical fields. Thus all three efforts showed that physical data assimilation increased surface phytoplankton concentrations, in places to the betterment of the model-data misfit, but the underlying cause was not investigated.

It is the purpose of this paper to examine in detail the impact of physical data assimilation on ecosystem dynamics. Our focus region is the California Current System (CCS), and the ecosystem model is a complex multi-component, self-assembling model (Follows et al., 2007). We apply a 4-Dimensional Variational Data Assimilation methodology that is increasingly common in regional analyses (Moore et al., 2011b; Matthews et al., 2012; Zhang et al., 2010). The paper is organized as follows. In section 'Methods', the model configuration for physical data assimilative and the ecosystem model are described briefly. Section 'Model results and comparison with observations' presents the ecosystem model results. A more complete discussion of the results and conclusions from the paper are given in section 'Discussion and conclusion'.

Methods

All modeling described in this study uses the Regional Ocean Modeling System (ROMS) (Schepetkin and McWilliams, 2005). The forward model configuration is similar to that discussed by Veneziani et al. (2009a,b). The model domain spans from the middle of the Baja peninsula to near the southern tip of Vancouver Island, and over 1000 km zonally, covering 30–48°N and 116–134°W, at $\frac{1}{10}^{\circ}$ resolution with 42 terrain-following vertical levels (Fig. 1). Surface forcing fields are derived from the high-resolution Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS; Doyle et al. (2009), provided by the Naval Research Laboratory). The daily COAMPS surface forcing data has a resolution of 3–9 km along the California and Oregon coasts. Boundary conditions are derived from global state estimates from the project, Estimating the Circulation and Climate of the Ocean (Wunsch and Heimbach, 2007; Wunsch et al., 2007) which contributed to the Global Ocean Data Assimilation Experiment.

The data assimilative framework is an incremental approach to variational assimilation (Courtier et al., 1994) that determines small adjustments in a control vector from prior values. The Incremental Strong Constraint Four-dimensional Variational Data Assimilation (I4D-Var) system has been incorporated recently into ROMS (Moore et al., 2011a). Sequential cycles assimilating sea surface height, sea surface temperature and available in situ hydrography have demonstrated lower root-mean squared errors in the Inter-America Seas (Powell et al., 2009), CCS (Broquet et al., 2009; Broquet et al., 2011), and Hawaiian waters (Matthews et al., 2012). The control vector in this study consists of the ocean state vector (horizontal velocity, temperature, salinity and sea surface elevation) at the start of the assimilation cycle along with the time-dependent wind stress, surface heat flux and fresh water flux. The horizontal decorrelation length scale was 300 km for wind stress and 100 km for heat and freshwater fluxes (Broquet et al., 2011). We used 14 day assimilation cycles based on analysis of the validity of the linearized dynamics of the tangent linear model (Veneziani et al., 2009b).

Surface temperature, sea surface height, and in situ observations of temperature and salinity are assimilated. The sea surface height observations are derived from the Aviso delayed-time sea level anomaly product and their estimate of Mean Dynamic Topography (Rio et al., 2005). The altimeter products were produced by Ssalto/Duacs and distributed by Aviso, with support from Centre National d' Etudes Spatiales (http://www.aviso.oceanobs.com/duacs/). We used a blended SST product, produced through CoastWatch/NOAA-Fisheries using the GEOS, AVHRR and MODIS platforms. This daily product represents a 5-day mean with a horizontal resolution of 0.1 degree and is available starting 22 July 2002. In situ T and S quality-controlled data were available from the UK Met Office (Ingleby and Huddleston, 2007) as part of the EU ENSEMBLE project (EN3). In the CCS, EN3 data includes CTD profiles that were collected during California Cooperative Oceanic Fisheries Investigation (CalCOFI; www.calcofi.org) and GLOBal ocean ECosystems dynamics (GLOBEC) cruises as well as data from the World Ocean Data 2005, and the Global Temperature Salinity Profile Program, Argo float data and XBT measurements.

The self-assembling ocean ecosystem model was developed by Follows et al. (2007) and its behavior in the CCS is described in Goebel et al. (2010). The NPZ-type model includes phytoplankton, zooplankton, dissolved and particulate organic matter, and multiple inorganic nutrients. This implementation includes 78 phytoplankton types, divided into four functional groups: diatoms, large non-diatoms, *Prochlorococcus*-like and small non-*Prochlorococcus*-like phytoplankton, each distinguished by their nutrient requirements. Diatoms require silicate for growth whereas non-diatoms do not. *Prochlorococcus*-like phytoplankton can take up ammonium and nitrite but not nitrate, whereas non-*Prochlorococcus*-like phytoplankton can utilize all three forms of inorganic nitrogen. Two size classes of zooplankton species are modeled, representing micro and mesozooplankton. Mesozooplankton graze on microzooplankton as well as phytoplankton.

The sequence of steps followed to obtain ecosystem model estimates for the forward model and I4D-Var is illustrated in Fig. 2. The physical model is integrated from climatological initial conditions for 6 years, driven by climatological surface forcing and lateral-boundary conditions. The coupled ecosystem model is then driven by realistic surface and boundary conditions from January 1, 1999 through July 22, 2002, using climatological initial conditions for nutrients, and small and uniformly distributed fields for other ecosystem variables. Within the model interior, model dynamics govern changes in fields (i.e., there is no relaxation to climatological values). Particulate material is allowed to sink out of the domain where it is lost to the system. A series of forward model cycles and I4D-Var data assimilative cycles proceeds until December 31, 2003. All results reported in this paper correspond to the model evaluation period from the January 1, 2003 through December 31, 2003.

As far as possible, identical model configurations and dynamical parameters were used for both the forward model and I4D-Var though some minor differences were unavoidable. Clamped boundary conditions were used for the assimilative run, but radiation boundary conditions (Marchesiello et al., 2001) were used for the non-data assimilative physical model. Generally, radiation boundary conditions result in fewer near boundary artifacts than clamped boundary conditions and thus are preferable. However, data assimilation suppresses these features to increase overall