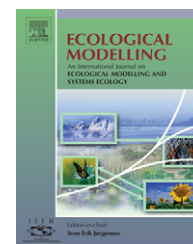


available at [www.sciencedirect.com](http://www.sciencedirect.com)journal homepage: [www.elsevier.com/locate/ecolmodel](http://www.elsevier.com/locate/ecolmodel)

# Calibration of the NEMURO nutrient–phytoplankton–zooplankton food web model to a coastal ecosystem: Evaluation of an automated calibration approach

Kenneth A. Rose<sup>a,\*</sup>, Bernard A. Megrey<sup>b</sup>, Francisco E. Werner<sup>c</sup>, Dan M. Ware<sup>d</sup>

<sup>a</sup> Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, LA 70803, USA

<sup>b</sup> National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sandpoint Way NE, Bin C15700, Seattle, WA 98115-0070, USA

<sup>c</sup> Department of Marine Sciences, University of North Carolina, Chapel Hill, NC 27599-3300, USA

<sup>d</sup> MRC, 3674 Planta Road, Nanaimo, BC V9T 1M2, Canada

## ARTICLE INFO

### Article history:

Published on line 9 November 2006

### Keywords:

Calibration  
Coastal  
Data assimilation  
Ecosystem model  
NEMURO  
North Pacific  
Nutrient–phytoplankton–zooplankton  
model  
Upwelling

## ABSTRACT

A one spatial-box version of the NEMURO oceanic lower trophic food web model was applied to a coastal upwelling environment typified by West Coast Vancouver Island. We used both ad hoc calibration and the automatic calibration program PEST. NEMURO was first calibrated to 1 year of monthly field data using the usual ad hoc approach of trial and error changes to 18 candidate parameters. Four PEST calibrations were then performed. The first three PEST calibrations used model predictions in year 10 from the ad hoc calibration as data in a twin experiment design; the fourth PEST calibration repeated the ad hoc procedure by having PEST calibrate NEMURO to the field data. When provided with ad hoc calibration model predictions as data, PEST accurately recovered the known 18 parameter values, even when small and large phytoplankton were lumped into total phytoplankton. When 57 parameters were allowed to vary PEST-estimated reasonable values for all 57 parameters, but they differed from the ad hoc calibrated values. However, when applied to the field data, PEST-estimated parameter values that differed greatly from the ad hoc values. The PEST calibration fitted some of the field data better than the ad hoc calibration but at the cost of unequal small and large phytoplankton concentrations. Thus, with proper and careful implementation, PEST offers a viable approach for objective calibration of NEMURO to site-specific monitoring data. We recommend that automatic calibration methods, such as PEST, be used for application of the NEMURO model to new locations. When the field data allow for specification of time series for each phytoplankton and zooplankton state variable, PEST will provide an objective, defensible, and repeatable way to calibrate the many parameters of the NEMURO model. If the available data are insufficient for specification of each state variable, then ad hoc calibration will likely be needed to allow for inclusion of qualitative decisions about model fit. Use of PEST in this situation will provide better understanding of the data-model mis-matches and will provide an alternative calibration to the necessary, but subjective, ad hoc calibration. Comparison of the ad hoc and PEST calibrations (even if unsuccessful) will help in the interpretation of the ad hoc calibration. Robust parameter estimation by any method depends on the quality and consistency of the calibration dataset.

© 2006 Elsevier B.V. All rights reserved.

\* Corresponding author. Tel.: +1 225 578 6346; fax: +1 225 578 6513.

E-mail address: [karose@lsu.edu](mailto:karose@lsu.edu) (K.A. Rose).

0304-3800/\$ – see front matter © 2006 Elsevier B.V. All rights reserved.

doi:10.1016/j.ecolmodel.2006.08.016

## 1. Introduction

Calibration of complex ecological models that have many potentially adjustable parameters has long been an issue in ecological modeling (Swartzman and Kaluzny, 1987). Most often an ad hoc calibration method is used whereby the modeler manually adjusts parameters until the fit between predicted and observed state variables are subsequently deemed “close enough” (Eckhardt and Arnold, 2001; Vrugt et al., 2003). However, one is never sure if the calibrated values are unique or optimal, and whether a different modeler would obtain the same values (Rose et al., 1991). One person’s good fit between predicted and observed values is often another person’s poor fit.

A variety of automatic methods for objective calibration of hydrological and water quality models have been proposed. Automatic, or mathematically formal calibration methods, are also sometimes referred to as data assimilation methods, although data assimilation also includes other variants such as sequential updating of model initial conditions and data smoothing and interpolation (Robinson and Lermusiaux, 2002). Examples of automatic calibration methods include the use of simulated annealing (e.g., Matear, 1995), genetic algorithms (e.g., Mulligan and Brown, 1998), the adjoint method (e.g., Schartau et al., 2001), and Bayesian-related methods (e.g., Vrugt et al., 2003). Vallino (2000) and Freedman et al. (1998) compared the performance of different optimization algorithms, while Madsen (2000) investigated multiple objective functions and Boyle et al. (2000) examined how to combine manual and automatic methods. These calibration methods share the general approach (Freedman et al., 1998; Vallino, 2000) of specifying an objective function based on data-model goodness-of-fit, and using accumulated information from previous model runs to determine how to change parameter values for subsequent runs. This is repeated until a minimum for the goodness-of-fit function is found; often checks, via using new starting parameter values or applying perturbations to parameter values, are then made to ensure the minimum is a global minimum of the objective function.

North Pacific Marine Ecosystem Model for Understanding Regional Oceanography (NEMURO) is a nutrient-phytoplankton-zooplankton food web model that was developed as part of an international effort (Werner et al., *this issue*), and is being applied to a variety of locations in the North Pacific. Examples of applications to date include NEMURO embedded in a one-dimensional vertical mixing model and applied to monitoring data at Station A-7 (Yamanaka et al., 2004) and Station KNOT (Fujii et al., 2002), imbedded in a three-dimensional ocean circulation model and examined at Stations P, A-7, and KNOT (Aita et al., 2003), used as one of four alternative models to simulate particle fluxes in the Western Pacific (Kishi et al., 2004), and used to assess the effects of vertical migration of zooplankton at Station A-7 (Yoshie et al., 2003) and the importance of the microbial food web off Hawaii (Smith et al., 2005). Herring and saury bioenergetics growth models have also been coupled to NEMURO (Megrey et al., *this issue*; Ito et al., 2004). Ongoing activities include further development of the biogeochemical processes represented in NEMURO, application

of NEMURO to other locations, and use of NEMURO coupled with bioenergetics models to examine climate effects on fish growth.

An important issue with the application of NEMURO, and other ecological models, is how to apply them in an objective way to different locations. Calibration is needed and appropriate, but how to calibrate the model remains an issue. Each modeler has their favorite parameters and opinions about the quality of the data and goodness of fit criteria (i.e., how close is close enough). Calibrating NEMURO to other locations is especially relevant as we pursue using NEMURO, as one mathematical expression of a nitrogen-phytoplankton-zooplankton model, to compare food web dynamics and fish growth among different geographic regions. Interpretation of model results as indicative of geographic differences in food web dynamics and fish growth requires that NEMURO be applied to the different regions in an objective and consistent manner. Otherwise, model predictions of geographic differences can be confounded with location-specific model calibration decisions.

The objective of this paper was to evaluate the utility and feasibility of using automatic calibration to apply NEMURO to a coast-like environment. Most NEMURO applications have involved open ocean locations. One direction in the continued development of NEMURO is to include herring that inhabit coastal areas. Megrey et al. (*this issue*) couple the version of NEMURO calibrated in this paper to a herring bioenergetics model. While Wainwright et al. (*this issue*) applied NEMURO, embedded in a two-dimensional spatial grid, to a coastal environment (California Current system), they used a version of NEMURO that used somewhat simplified kinetics and constant water temperature. We needed a version of NEMURO that could be used as part of analysis of herring growth for a coastal upwelling system. We used data from West Coast of Vancouver Island (WCVI) as the basis for calibrating NEMURO to a coastal system. WCVI is a good candidate for evaluating calibration methods because of the extensive field data on the lower trophic levels (Robinson and Ware, 1999; Mackas et al., 2004) and because we knew the next step was to attempt to simulate herring dynamics in that region (Megrey et al., *this issue*).

In this paper, we calibrate NEMURO to a coastal environment using the usual ad hoc method (manual adjustment of parameter values) and the PEST (Doherty, 2004) automated calibration method. The PEST method uses similar algorithms as other commonly used methods (e.g., variational adjoint technique—Friedrichs, 2002), but has the added advantage of being able to be used with an external model code thereby reducing the need for new coding. We first briefly describe the NEMURO model, and how we modified NEMURO for a coastal upwelling system. We then describe and present the results of an ad hoc calibration and four PEST calibrations, all based on 1 year of monthly phytoplankton and zooplankton data. All calibrations started with the same set of parameter values. The first three PEST calibrations used the ad hoc calibration results as observed data and were designed to determine if PEST could estimate known parameter values, the effect of lumping state variables, and the effect of allowing many parameters to be varied. Use of the same model to generate the data and then recover parameter values from the generated data is sometimes referred to as an “identical twin experi-

ment” (Robinson and Lermusiaux, 2002). Identical twin experiments are commonly used to evaluate automatic calibration methods (e.g., Harmon and Challenor, 1997). The fourth PEST calibration used the assembled field data as observed data and was designed to see if PEST could successfully calibrate NEMURO to our coastal field dataset, and if so, how PEST-estimated parameter values compared with our ad hoc estimated values. We conclude with a discussion of the lessons learned from our calibration effort, and the general merits of using automated parameter estimation methods such as PEST for calibrating ecological models for regional comparisons.

## 2. Description of NEMURO

NEMURO implemented in this paper simulates the dynamics of the nutrient–phytoplankton–zooplankton food web in a single well-mixed spatial box that represents the surface layer of the water column (Fig. 1; Kishi et al., 2007). The food web is represented with eleven state variables: nitrate ( $\text{NO}_3$ ), ammonium ( $\text{NH}_4$ ), small phytoplankton (PS), large phytoplankton (PL), small zooplankton (ZS), large zooplankton (ZL), predatory zooplankton (ZP), particulate organic nitrogen (PON), dissolved organic nitrogen (DON), particulate organic silicon (Opal), and silicic acid ( $\text{Si}(\text{OH})_4$ ). All state variables are tracked in the units of  $(\mu\text{mol N}) \text{L}^{-1}$ ; silicon variables are converted from nitrogen using constant silicon to nitrogen ratios. NEMURO is a system of 11 coupled ordinary differential equations, with an equation for the rate of change of each state variable. NEMURO was solved using a fourth-order Runge–Kutta solution technique

with an integration time step of 1 h. The fluxes between the state variables are shown in Fig. 1 for nitrogen (solid arrows) and silicon (dashed arrows). The driving variables are annual cycles of daily water temperature, daily incident solar radiation, and the mixed layer depth (and associated fluxes of nutrients into the modeled box).

The rate of change of each state variable is expressed as the sum of process rates that affect that state variable. Photosynthesis, respiration, excretion, predation by zooplankton in the model and other (not accounted for by simulated zooplankton) mortality affect each phytoplankton state variable; grazing, egestion, excretion, predation by zooplankton, and other mortality affect each zooplankton state variable. Nutrient state variables are reduced by photosynthesis uptake, increased by various combinations of phytoplankton and zooplankton respiration, excretion, mortality, and converted among nutrient forms via first-order, temperature-dependent decomposition reactions. Phytoplankton photosynthesis, respiration, and mortality, and zooplankton grazing and other mortality, are all temperature-dependent. Parameters are expressed as rates at  $0^\circ\text{C}$ , and  $Q_{10}$  functions are used to adjust these rates each day based on the simulated water temperature. Photosynthesis also depends on the average light (integrated over the water column) and nitrate and ammonium concentrations, with photosynthesis of large phytoplankton (diatom-like) also dependent on silicic acid. Grazing by zooplankton is dependent on the concentrations of prey using a formulation in which an Ivlev parameter governs how quickly maximum grazing rates are approached with increasing prey concentrations.

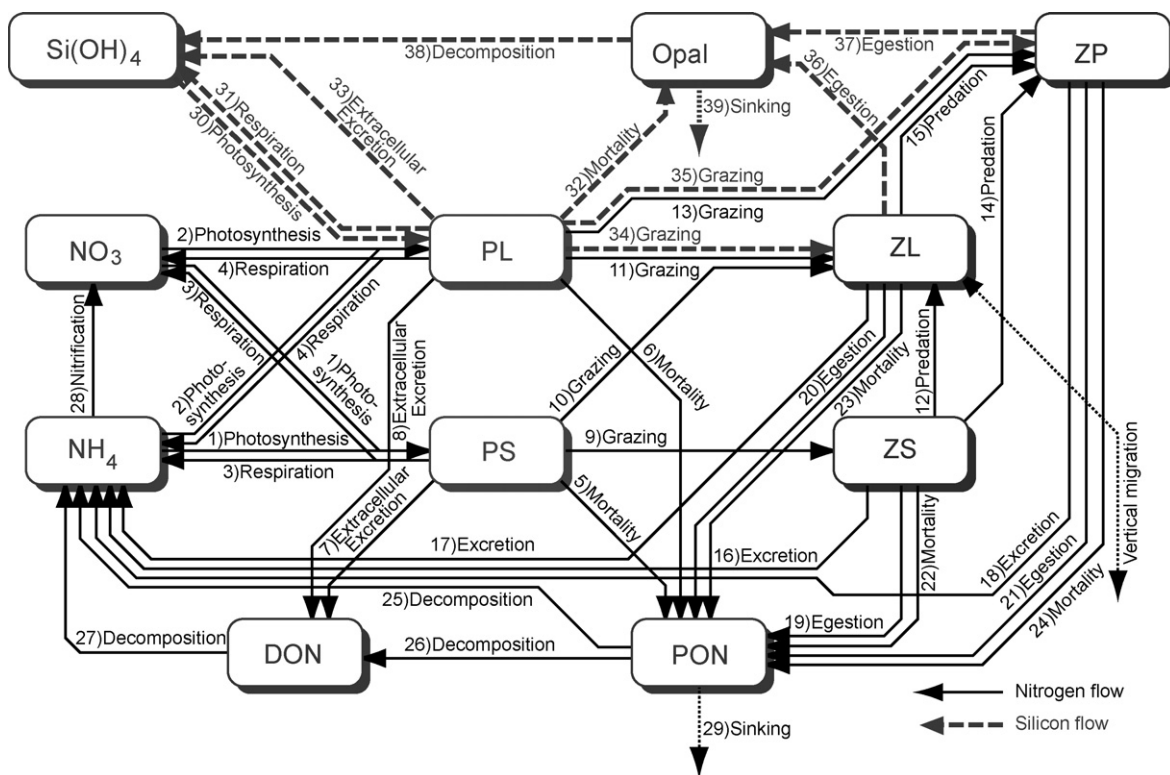
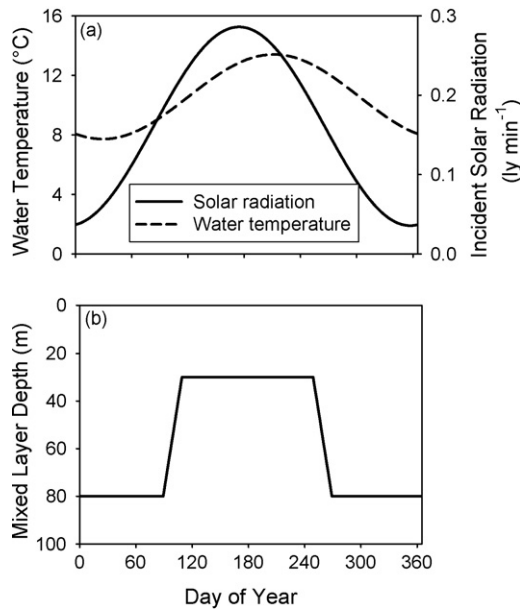


Fig. 1 – Food web diagram of the 11 state variables represented in the NEMURO model showing the key processes and the flows of nitrogen and silicon (reproduced from Kishi et al., this issue). Vertical migration of the large zooplankton was eliminated for application in this paper.



**Fig. 2 – Driving variables used as representative of the coastal waters of West Coast Vancouver Island (WCVI) area. (a) Daily water temperature and incident solar radiation and (b) mixed layer depth.**

We modified the driving (abiotic) variables of NEMURO to crudely represent a coastal upwelling system, and we eliminated the seasonal vertical migration of the large zooplankton. We formulated functions that generated daily water temperatures and incident solar radiation representative of the WCVI area (Fig. 2a). We modified the open-ocean formulation for the mixed layer depth and nutrient fluxing from below from previous applications to roughly approximate an upwelling system. Mixed layer depth was set to 80 m for 26 September–30 March, and to 30 m for 9 April–6 September (Robinson et al., 1993), with linear interpolation over 10 days for each of the transitions (Fig. 2b). The influence of the seasonal dynamics of the mixed layer depth on NEMURO predictions deserves additional investigation as recent information on mixed layer depth suggest shallower depths may be appropriate (Thomson and Fine, 2003). Flux of nitrate into the modeled box was computed as the difference between the concentration in the modeled box and  $25.0 \mu\text{mol NL}^{-1}$ , times an exchange rate. We coupled the nitrate flux into the modeled box to the changes in the mixed layer depth by assuming the computed flux when the mixed layer depth was 80 m, and five times the computed flux whenever the mixed layer depth was less than 80 m. Because of the shallow waters in this coastal environment, we eliminated the seasonal vertical migration of the large zooplankton. Large zooplankton is always present in the modeled spatial box.

### 3. Coastal field data

We opted to calibrate NEMURO to an idealized dataset representative of a coast-like environment similar to the WCVI. We recognize that representing a coastal upwelling system with a

single box model is difficult and unrealistic. A spatially explicit approach that allows for inclusion of physical features critical to realistic simulation of an upwelling system is presented by Wainwright et al. (this issue). We were interested in generating zooplankton densities of similar magnitude and seasonal patterns as a coastal upwelling system to use as prey for a herring growth model described in Megrey et al. (this issue).

We used field data from WCVI and California Current system to construct monthly values of total phytoplankton, small zooplankton, large zooplankton, and predatory zooplankton concentrations as the basis for calibration. These assembled field data were from a variety of sources and time periods so we constructed an idealized annual cycle of monthly values. Our idealized dataset is not intended to represent a specific year or an exact location. Only WCVI field data for monthly concentrations were used in model calibration.

The primary calibration dataset consisted of multiple years (1991–2001) of measured concentrations of total phytoplankton (chlorophyll), small copepods, medium copepods, large copepods, total zooplankton, and euphasiids for several locations off WCVI. The data were assembled from a variety of sources: chlorophyll data from Richard Thomson at the Institute of Ocean Sciences (Fisheries and Oceans Canada, Sidney, BC), and zooplankton data from Stephen Romaine via the Institute of Ocean Sciences' Zooplankton Database and Donald McQueen at Aquatic Ecosystem Associates, Nanaimo, BC. Our zooplankton data includes much of the zooplankton data reported by Mackas (1992).

We assumed euphasiids in the monitoring data equated to predatory zooplankton in NEMURO, and we then calculated the concentration of large zooplankton in NEMURO as total zooplankton minus euphasiids. Because of the mesh size used in the zooplankton sampling, the WCVI monitoring data included zooplankton generally greater than  $200 \mu\text{m}$  (i.e., not microzooplankton). We refer to total zooplankton from monitoring data as total zooplankton greater than 200 microns and equate total zooplankton greater than  $200 \mu\text{m}$  with large plus predatory zooplankton in NEMURO.

Estimation of monthly small zooplankton concentrations was achieved by use of additional data reported for the California Current. Wang (1998) reported concentrations of nauplii, copepodites, and adult copepods for the California Current. We computed monthly average concentrations of each of these, and then computed the average ratio of nauplii to copepodites plus adult copepods. We then applied this ratio to the computed total copepods (small plus large) in the WCVI dataset to obtain the monthly concentrations of small zooplankton for NEMURO. We assumed that copepodites plus adult copepods in Wang's California Current data was similar to the summed (greater than  $200 \mu\text{m}$ ) copepods in the WCVI data. Our estimates of small and predatory zooplankton concentrations are likely low because we only accounted for nauplii in the small zooplankton (i.e., did not include heterotrophic flagellates and ciliates) and we did not explicitly account for chaetognaths (Mackas, 1992) in the predatory zooplankton.

We used a third field dataset reported in Wainwright et al. (this issue) for the California Current system as a secondary source of information for model calibration. Wainwright et al. (this issue) reported nitrate, total phytoplankton (chlorophyll),



and total zooplankton for three different depths off of the Oregon coast. These data were not used in the automatic calibration but rather were used to provide some qualitative confirmation that the idealized WCVI monthly data were realistic.

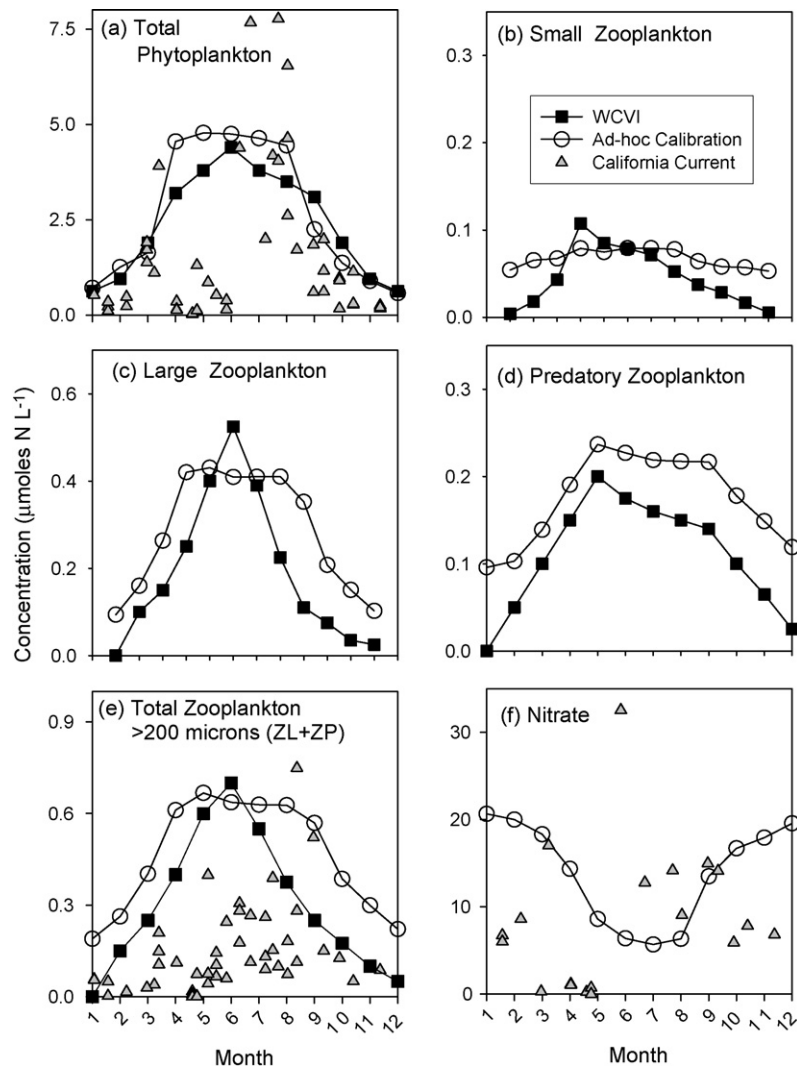
The idealized monthly values in our calibration dataset showed a broad peak of total phytoplankton (solid square symbols in Fig. 3a), and showed zooplankton dominated by the large zooplankton group (solid square symbols in Fig. 3c). Monthly total phytoplankton and total zooplankton in the calibration dataset were consistent with the concentrations reported by Wainwright et al. (this issue) for the California Current system (grey triangles in Fig. 3a and e).

#### 4. Calibration methods

We used the ad hoc method and PEST to calibrate NEMURO to the idealized coastal dataset. Ad hoc calibration was the

usual manual adjustment of parameters values by one of the authors (KAR) until we decided that plots of predicted monthly total phytoplankton and zooplankton densities over 1 year closely enough mimicked the observed values.

The second method was the use of an automated calibration method called PEST (Doherty, 2004). PEST uses a variation of the Gauss–Marquardt–Levenberg algorithm to determine the values of parameters that minimize the weighted sum of squared deviations between predicted and observed values. PEST approximates the relationship between observations and model parameters using a Taylor series expansion, which involves the Jacobian matrix (the matrix of partial derivatives of observations with respect with parameters). PEST computes the partial derivatives of the Jacobian matrix using central differences; parameters are varied small amounts from the current set of values, the model is rerun, and the derivatives computed. New values of parameters are determined (parameter update vector) based on using the Marquardt



**Fig. 3** – Idealized coastal dataset for WCVI (solid square), field data reported for the California Current (grey triangles), and model predictions from the ad hoc calibration (open circle). WCVI calibration data was not available for nitrate. California Current data was available for total phytoplankton, total zooplankton (greater than 200 µm), and nitrate. (a) Total phytoplankton, (b) small zooplankton, (c) large zooplankton, (d) predatory zooplankton, (e) total zooplankton greater than 200 µm (ZL + ZP from NEMURO), and (f) nitrate.

parameter (denoted  $\lambda$ ), which is based on the gradient of the objective function (derivative of the objective function with respect to parameters). PEST stops searching when the objective function does not go lower over several iterations, when the changes in parameters dictated by the update vector are very small, or when the number of iterations or other internal calculations are triggered. Published examples of the use of PEST for model calibration involve a water shed scale model of water flow (Al-Abed and Whiteley, 2002; Doherty and Johnston, 2003), a model of nutrient export from a stream (Baginska et al., 2003), and radionuclide transport in groundwater models (Zyvoloski et al., 2003).

4.1. Model simulations and starting parameter values

All model simulations were for 12 years, with the annual cycle of daily driving variables (mixed layer depth, nitrate influx, water temperature, and incident solar radiation) repeated year after year. Interannual differences in model predictions were relatively small compared to the seasonal differences within each year. We focused our calibration on monthly values during year 10 of the 12-year simulations, although any of the years would have yielded similar results. We also examined the dynamics over the 12 years to ensure that the calibrated model was stable for multiple years, and that differences among alternative calibrations were maintained over years (i.e., differences were not just transient).

All calibration attempts started with a set of NEMURO parameter values generally consistent with values used in previous applications. We wanted a set of starting parameter values that generated phytoplankton and zooplankton concentrations different from our idealized coastal dataset (Fig. 3) in order to set-up a challenging calibration problem. We used a recent set of parameter values that resulted from NEMURO being applied to monitoring data from Station A-7 (Yamanaka et al., 2004). We changed the temperature effects parameters for small and large phytoplankton respiration from 0.0519 ( $Q_{10} = 1.68$ ) used by Yamanaka et al. (2004) to 0.0693 ( $Q_{10} = 2.0$ ) to match the temperature effects parameter values used for maximum photosynthetic rate and mortality rate. We refer to these parameter values as the starting values, and a model simulation using these values differed from the idealized monthly field data (see Fig. 4).

4.2. Ad hoc calibration

We calibrated NEMURO by changing key model parameters in an attempt to match the monthly total phytoplankton and zooplankton concentrations time series of our coastal dataset. We began with all model parameters at their starting values. We restricted our adjustments to what we deemed 18 candidate parameters, keeping all of the other parameters at their starting values. The eighteen parameters we varied were the maximum photosynthetic rates of small and large phytoplankton, the seven maximum grazing rates of the zooplankton, the three Ivlev grazing parameters, and the mortality rates and temperature effect parameters of the zooplankton (Table 1). We selected these parameters based on our previous experience with NEMURO. Yoshie et al. (2007) conducted a sensitivity analysis of a one-box version of NEMURO applied to

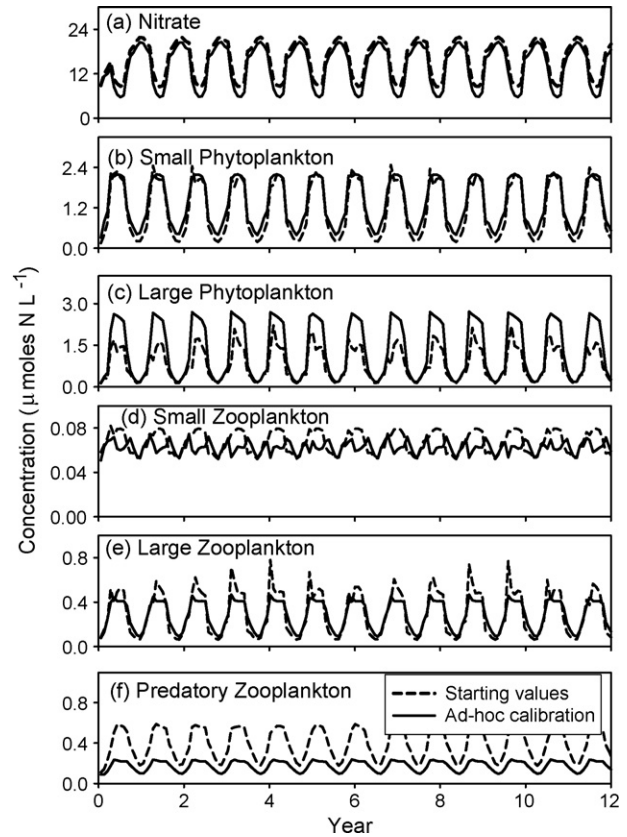


Fig. 4 – Predicted nitrate, phytoplankton, and zooplankton for 12-year simulations using the starting values of parameters and the ad hoc calibrated values of parameters.

Stations A-7 and P. They concluded that maximum photosynthesis rates, maximum grazing rates, and the Ivlev parameter and mortality-related parameters for predatory zooplankton were important. Calibration consisted of adjusting any of these eighteen parameters until we could get predicted monthly values to mimic what we deemed important patterns in our calibration dataset.

4.3. PEST calibrations

We performed four automated parameter calibrations using the PEST software (Table 2). The first calibration was designed to test whether PEST could estimate parameter values when we know the parameter values that generated the data (i.e., identical twin experiment). We used the predicted monthly values of small and large phytoplankton, and small, large, and predatory zooplankton from year 10 of the ad hoc calibrated simulation as observed data for PEST (Table 3). We began PEST with our starting values and allowed PEST to adjust the 18 candidate parameters to try to match the ad hoc calibrated values of phytoplankton and zooplankton. We compared parameter values and NEMURO predictions between the ad hoc calibration and the PEST calibration.

The second PEST calibration was designed to confirm that PEST could estimate known parameter values with the small and large phytoplankton lumped into a single prediction variable. Our calibration dataset synthesized from the monitoring

**Table 1 – The 18 candidate parameters of the NEMURO model varied in the ad hoc calibration and three of the four PEST calibration simulations**

Parameter	Definition	Units
Vmaxs	Maximum photosynthetic rate at 0 °C of small phytoplankton	d <sup>-1</sup>
Vmaxl	Maximum photosynthetic rate at 0 °C of large phytoplankton	d <sup>-1</sup>
Grmaxs	Maximum grazing rate at 0 °C of small zooplankton on small phytoplankton	d <sup>-1</sup>
Lams	Ivlev constant governing grazing rate of small zooplankton	μmol N <sup>-1</sup>
Morzso	Rate constant at 0 °C of small zooplankton for quadratic mortality rate relationship	L (μmol N) <sup>-1</sup> d <sup>-1</sup>
Kmorzs	Temperature effect on mortality rate of small zooplankton	°C <sup>-1</sup>
Grmaxlps	Maximum grazing rate at 0 °C of large zooplankton on small phytoplankton	d <sup>-1</sup>
Grmaxlpl	Maximum grazing rate at 0 °C of large zooplankton on large phytoplankton	d <sup>-1</sup>
Grmaxlzs	Maximum grazing rate at 0 °C of large zooplankton on small zooplankton	d <sup>-1</sup>
Laml	Ivlev constant governing grazing rate of large zooplankton	L (μmol N) <sup>-1</sup>
Morzlo	Rate constant at 0 °C of large zooplankton for quadratic mortality rate relationship	L (μmol N) <sup>-1</sup> d <sup>-1</sup>
Kmorzl	Temperature effect on mortality rate of large zooplankton	°C <sup>-1</sup>
Grmaxppl	Maximum grazing rate at 0 °C of predatory zooplankton on large phytoplankton	d <sup>-1</sup>
Grmaxpzs	Maximum grazing rate at 0 °C of predatory zooplankton on small zooplankton	d <sup>-1</sup>
Grmaxpzl	Maximum grazing rate at 0 °C of predatory zooplankton on large zooplankton	d <sup>-1</sup>
Lamp	Ivlev constant governing grazing rate of predatory zooplankton	L (μmol N) <sup>-1</sup>
Morzpo	Rate constant at 0 °C of predatory zooplankton for quadratic mortality rate relationship	L (μmol N) <sup>-1</sup> d <sup>-1</sup>
Kmorzp	Temperature effect on mortality rate of predatory zooplankton	°C <sup>-1</sup>

**Table 2 – Characteristics of the four PEST calibration simulations**

PEST simulation	Source of observed data	No. of parameters allowed to vary	Small and large phytoplankton
1	Ad hoc calibration	18	Separate
2	Ad hoc calibration	18	Combined
3	Ad hoc calibration	57	Separate
4	Idealized field dataset	18	Combined

The source of observed data used by PEST was either the output from the ad hoc simulation or the idealized coastal dataset. The number of parameters varied was either the 18 candidate parameters (Table 1) or 57 parameters related to phytoplankton and zooplankton. Small and large phytoplankton groups were either treated as separate prediction variables or summed as a single variable of total phytoplankton. Constraints on parameters were roughly order of magnitude (e.g., 0.01 and 10.0 for a value of 0.4) for all four PEST simulations.

data had chlorophyll as the measure of total phytoplankton. We repeated the first PEST calibration, including using the Station A-7 calibrated values as starting values, but provided PEST monthly values of small plus large phytoplankton, rather than small and large phytoplankton as separate variables.

The third PEST calibration was designed to see the effects of allowing PEST to vary more than the 18 candidate parameters. We allowed 57 parameters, including the 18 candidate parameters, related to phytoplankton and zooplankton to be varied by PEST.

The fourth calibration was the application of PEST to the field-estimated calibration dataset. Given the results of PEST calibrations 1–3 showed that PEST could recover known parameter values under separate and lumped small and large phytoplankton and when various numbers of parameters were allowed to be varied, we challenged PEST with the same dataset we used as our starting point for the ad hoc calibration. We restricted PEST to calibrate the same 18 candidate parameters we varied in the ad hoc calibration. We wanted to see whether PEST, when presented with the same calibration

**Table 3 – Monthly concentrations (μmol NL<sup>-1</sup>) of phytoplankton and zooplankton from year 10 of the ad hoc calibration simulation used as observed data for PEST calibration simulations 1–3**

Month	Small phytoplankton	Large phytoplankton	Small zooplankton	Large zooplankton	Predatory zooplankton
January	0.510	0.206	0.054	0.094	0.096
February	0.837	0.426	0.065	0.160	0.103
March	1.058	0.585	0.067	0.263	0.139
April	2.150	2.403	0.079	0.420	0.191
May	2.145	2.634	0.075	0.431	0.237
June	2.196	2.554	0.079	0.409	0.227
July	2.174	2.465	0.079	0.410	0.219
August	2.097	2.351	0.078	0.410	0.217
September	1.081	1.169	0.064	0.352	0.217
October	0.807	0.564	0.058	0.208	0.178
November	0.590	0.308	0.057	0.150	0.149
December	0.405	0.173	0.053	0.102	0.119

situation as we faced in our ad hoc calibration, would generate similar model fits to the data with similar parameter values.

As with the ad hoc calibration, all PEST calibrations used 12-year simulations and based its goodness of fit on predicted and observed monthly values for year 10. All PEST calibrations started with our starting parameter values. We used log transformation option on all adjustable parameters (see Doherty, 2004), which allows for relatively large changes in parameter values that helps the efficiency of the optimization. Minimum and maximum values for parameters were specified as roughly an order of magnitude in either direction. For example, a parameter value of 0.4 would have minimum value of 0.01 and a maximum value of 10.0.

We summarize the fit to the calibration data (either the ad hoc model output or the field data) with the average of the residuals (data minus predicted) and the largest (in absolute value) individual residual. The averaged residuals relate to the overall bias associated with model predictions, while the largest absolute residual provides information on the general magnitude of the residuals to ensure that small bias is not due to the canceling effects of very large positive and negative residuals. For PEST simulations 1–3, predicted daily concentrations of phytoplankton and zooplankton concentrations over the 12 years were virtually identical to concentrations from the ad hoc simulation. We simply note this in the results and refer the reader to the ad hoc predictions shown in Fig. 4. Throughout the text, we refer to parameters by their names, which are defined in Table 1.

## 5. Results

### 5.1. Ad hoc calibration

We were able to calibrate NEMURO to the idealized coastal dataset by varying some of the 18 candidate parameters within

reasonable ranges of values (Table 4). Once the final ad hoc calibration was achieved, it turned out that four of the candidate parameters (two Ivlev grazing parameters: Laml and Lamp; two of the zooplankton maximum grazing rates: Grmaxlps and Grmaxpzl) had not been changed from their starting values. Predicted total phytoplankton, and predicted small, large, and predatory zooplankton concentrations were similar in magnitude and exhibited similar seasonal patterns to the calibration data (open circles versus solid squares in Fig. 3). The model-calibrated peak of small zooplankton was flatter and did not go as close to zero as the field data indicated (Fig. 3b). This was deemed acceptable because we know from past experience that the NEMURO model has difficulty generating very small starting and ending values of biomass for the small and predatory zooplankton state variables. Predicted nitrate concentration was similar in magnitude to nitrate concentrations reported in the secondary dataset from the California Current system (Fig. 3f). Once calibration was concluded, we computed the averaged residual as  $-0.11 \mu\text{mol NL}^{-1}$ , with the largest individual residual value of  $-1.35 \mu\text{mol NL}^{-1}$  for total phytoplankton in April.

Relatively large adjustments were made to some of the 18 candidate parameters that were allowed to be varied. For example, the temperature effects on grazing parameters for the small, large, and predatory zooplankton were adjusted from its starting value of 0.0693 ( $Q_{10} = 2.0$ ) to 0.1099 ( $Q_{10} = 3.0$ ), and some of the zooplankton maximum grazing rates were significantly reduced, with the largest reduction being for large zooplankton eating small zooplankton (Grmaxlzs: starting value of 0.40–0.14  $\text{d}^{-1}$ ).

Twelve-year simulations using starting values of parameters and ad hoc calibrated values of parameters illustrate the magnitude of the calibration challenge (Fig. 4). Peak concentrations of calibrated large phytoplankton (Fig. 4c) were almost twice the peak concentrations based on starting value parameters, while peak concentrations of large and predatory

**Table 4 – Starting values and the estimated values of the 18 candidate parameters for the ad hoc simulation and the five PEST calibration simulations**

Parameter	Starting values	Ad hoc	PEST 1	PEST 2	PEST 3	PEST 4
Vmaxs	0.4	0.49	0.49	0.49	0.52	0.56
Vmaxl	0.8	0.71	0.71	0.71	0.81	0.25
Grmaxs	0.4	0.31	0.31	0.32	0.33	0.086 <sup>a</sup>
Lams	1.4	0.4	0.412	0.404	0.404	0.497
Morsz0	0.677	0.877	0.886	0.884	0.889	0.943
Kmorzs	0.0693	0.1099	0.1099	0.1081	0.0905	0.0902
Grmaxlps	0.1	0.1	0.1	0.1	0.13	0.28
Grmaxlpl	0.4	0.23	0.23	0.23	0.27	0.034
Grmaxlzs	0.4	0.14	0.14	0.14	0.26	0.86 <sup>a</sup>
Laml	1.4	1.4	1.4	1.4	1.38	0.1
Morzl0	0.677	0.877	0.877	0.875	1.035	0.196
Kmorzl	0.0693	0.1098	0.1098	0.1099	0.0866	0.0294
Grmaxppl	0.2	0.11	0.11	0.11	0.15	0.054
Grmaxpzs	0.2	0.11	0.11	0.12	0.17	0.21
Grmaxpzl	0.2	0.20	0.20	0.20	0.25	0.15
Lamp	1.4	1.4	1.4	1.4	1.216	1.16
Morzp0	0.677	0.877	0.877	0.879	0.977	1.183
Kmorzp	0.0693	0.1099	0.1099	0.1098	0.0977	0.0490

<sup>a</sup> Calibrated value of the parameter is at minimum or maximum allowed value.



zooplankton (Fig. 4e and f) were roughly half the peak concentrations based on starting value parameters.

### 5.2. PEST simulation 1

When presented with model predictions as observed data, PEST-estimated values of the 18 candidate parameters that were practically identical to the ad hoc calibrated values (Table 4, PEST 1 versus ad hoc). The only differences were for the Ivlev parameter for small zooplankton (Lams: 0.412 in PEST versus 0.4 in ad hoc), and the mortality parameter for small zooplankton (Morsz0: 0.886 in PEST versus 0.877 in ad hoc). All residuals were very small and averaged  $1.22 \times 10^{-6}$  with a largest single residual of  $3.0 \times 10^{-5}$  for small phytoplankton biomass in October. Although PEST was given the opportunity to change all 18 parameters, PEST correctly retained the starting values of the four parameters unchanged during ad hoc calibration. Given that most PEST-obtained parameters values were nearly identical to ad hoc-obtained values, predicted concentrations of nitrate, phytoplankton, and zooplankton with the PEST-estimated values were indistinguishable from those using the ad hoc calibrated values shown in Fig. 4.

### 5.3. PEST simulation 2

PEST successfully estimated the known parameter values with small and large phytoplankton combined into a single variable of total phytoplankton (Table 4, PEST 2 versus ad hoc). Nine of the parameters estimated by PEST had identical values to the ad hoc calibrated values, and another seven parameters only differed in their last digit (e.g., maximum grazing rate of small zooplankton on small phytoplankton (Grmaxs) was 0.32 by PEST versus 0.31 in ad hoc). The remaining two parameters related to mortality of small zooplankton differed the most between PEST and ad hoc calibrated values (Morzs0: 0.884 by PEST versus 0.877 in ad hoc; Kmorzs: 0.1081 by PEST versus 0.1099 in ad hoc). Residuals were all small and averaged  $3.89 \times 10^{-6}$ , with the largest residual being  $1.1 \times 10^{-4}$  for predatory zooplankton in October. Predicted phytoplankton and zooplankton using the PEST-estimated parameter values were again virtually identical to the predicted concentrations using the ad hoc calibrated parameter values shown in Fig. 4.

### 5.4. PEST simulation 3

Increasing the number of parameters that could be varied to 57 affected how well PEST could estimate the known values of the 18 candidate parameters (Table 4). The averaged and largest residuals were  $5.63 \times 10^{-5}$  and  $-1.07 \times 10^{-3}$  for small zooplankton biomass in October, with PEST-predicted phytoplankton and zooplankton concentrations virtually identical to the ad hoc calibrated concentrations shown in Fig. 4.

Ideally, in the extreme case of allowing 57 parameters to be varied, PEST would have estimated the same values for the 14 of 18 candidate parameters known to have been varied in the ad hoc calibration. When we allowed 57 parameters to be varied, PEST generated reasonable values for all 57 parameters, including the 18 candidate parameters. The largest differences among the candidate parameters between

PEST and ad hoc values were: 0.26 versus 0.14 for Grmaxlzs and 1.035 versus 0.877 for Morzl0 (Table 4, PEST 3 versus ad hoc). PEST correctly estimated values close to the ad hoc values for the four candidate parameter not changed during the ad hoc calibration (Grmaxlps, Lam1, Grmaxpzl, and Lamp). PEST-calibrated values of the remaining candidate parameters were similar, but not identical, to their ad hoc calibrated values (Table 4).

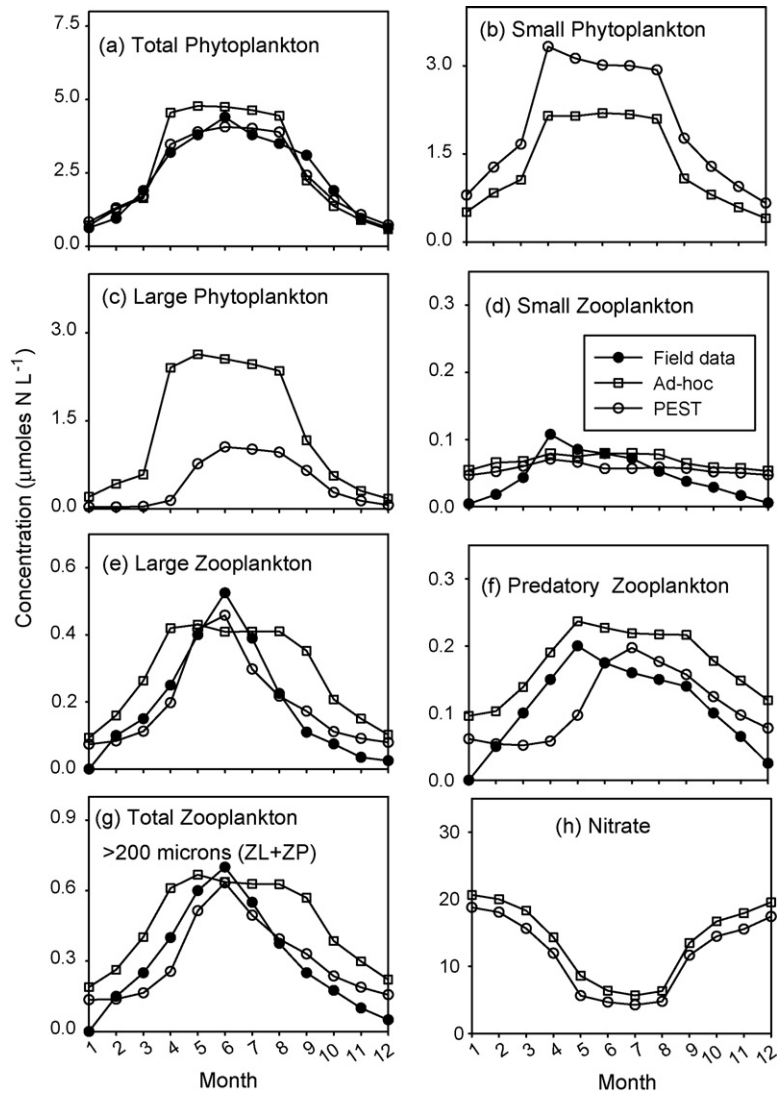
### 5.5. PEST simulation 4

When confronted with the field dataset, PEST was marginally successful in estimating values for the 18 candidate parameters that generated a reasonably close fit to the field data (open circle versus close circle in Fig. 5). The PEST calibration generated total phytoplankton (Fig. 5a) and large zooplankton (Fig. 5e) concentrations that were similar to the field dataset; predatory zooplankton concentration were similar but with a delayed peak (Fig. 5f). As with the ad hoc calibration, PEST had difficulty getting small zooplankton to exhibit sharp enough peaks and low enough troughs (Fig. 5d).

PEST-generated phytoplankton and zooplankton concentrations appeared to be similar, or more similar, to the field data as the ad hoc predicted concentrations (open circle versus open square in Fig. 5). The averaged residual was  $-0.0081$  for the PEST calibration, which was lower than the averaged residual of  $-0.11$  obtained for the ad hoc calibration; the single largest residual was 0.68 for PEST (total phytoplankton in September) and 1.35 for the ad hoc simulation (total phytoplankton in April). PEST was better able to match total phytoplankton concentrations (Fig. 5a), large zooplankton concentrations (Fig. 5e), and total zooplankton concentrations (Fig. 5g) than the ad hoc calibration. However, PEST achieved its fit to total phytoplankton with higher small phytoplankton and lower large phytoplankton than the ad hoc calibration (Fig. 5b and c). Both PEST and ad hoc calibration had difficulties with the peaks and troughs of the small zooplankton (Fig. 5d) and predatory zooplankton (Fig. 5f). PEST predicted reasonably close peaks and troughs of predatory zooplankton but the predicted peak was delayed a few months compared to the field data. The ad hoc calibration predicted the peak concentration during the correct month but consistently overestimated predatory zooplankton concentrations. Predicted nitrate concentrations were similar between the PEST and ad hoc calibrations (Fig. 5h).

Most of the values of the eighteen candidate parameters estimated by PEST differed from those determined by ad hoc calibration (Table 4, PEST 4 versus ad hoc). For example, maximum grazing rates of large zooplankton estimated by PEST versus ad hoc calibration were: 0.28 versus 0.1 for Grmaxlps, 0.034 versus 0.23 for Grmaxlpl, and 0.86 versus 0.14 for Grmaxlzs. Two parameters (Grmaxs and Grmaxlzs) were estimated by PEST to have values at the minimum values we allowed.

The differences in the fit of PEST and ad hoc calibration to the 1 year of the field dataset (Fig. 5) manifested themselves in the long-term simulations (Fig. 6). PEST generated higher peak concentrations of small phytoplankton (Fig. 6b) and lower peak concentrations of large phytoplankton (Fig. 6c), and slightly lower large zooplankton (Fig. 6e) and predatory



**Fig. 5 – Idealized coastal data for WCVI (solid circle), and model predictions from the ad hoc calibration (open square) and PEST calibration with PEST applied to the field data (open circle). WCVI calibration data was not available for small phytoplankton, large phytoplankton, and nitrate. (a) Total phytoplankton, (b) small phytoplankton, (c) large phytoplankton, (d) small zooplankton, (e) large zooplankton, (f) predatory zooplankton, (g) total zooplankton greater than 200 µm (ZL + ZP from NEMURO), and (h) nitrate.**

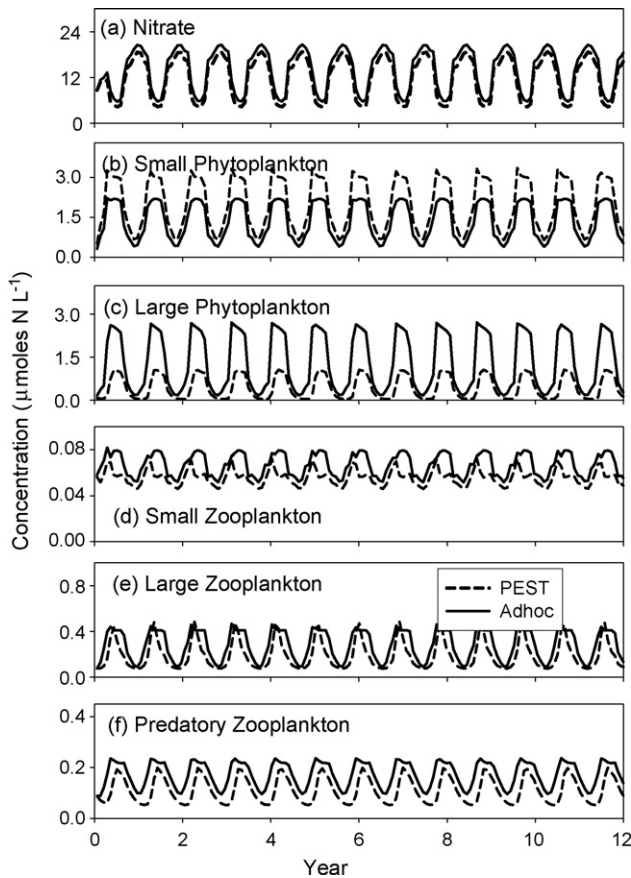
zooplankton concentrations (Fig. 6f). PEST generated small zooplankton concentrations that were lower and peaked earlier in the year than the ad hoc calibrated values (Fig. 6d).

## 6. Discussion

Our situation is common to many modeling applications. We wanted to apply a model to new locations for which we had limited data assembled from a variety of sources. In order for us to interpret model prediction differences between locations as truly location differences, we wanted to devise a consistent scheme for model calibration. While we will never eliminate modeler decisions involved with new model applications, we wanted to minimize and document these decisions as much as possible. Automatic calibration methods can help eliminate

one source of arbitrariness in modeler decisions by standardizing how the model is calibrated to monitoring data at the different locations.

Our analyses using the PEST software to objectively calibrate NEMURO are encouraging. Under constrained conditions, PEST-estimated known parameter values. PEST simulations 1 and 2 were constrained and set-up almost ideal conditions for automatic calibration by PEST. In PEST simulations 1 and 2 we only allowed PEST to vary the 18 parameters we knew that we had considered varying in the ad hoc calibration. For small and large phytoplankton treated as two separate groups (PEST 1) and combined into a single total phytoplankton variable (PEST 2), PEST was able to estimate parameter values virtually identical to the true ad hoc values (Table 4). When we increased the number of parameters allowed to be varied by PEST to 57 in simulation



**Fig. 6 – Predicted nitrate, phytoplankton, and zooplankton for 12-year simulations using the ad hoc calibrated values of parameters and PEST-estimated values with PEST applied to the field data.**

3, estimated values of the eighteen candidate parameters diverged from the ad hoc values but were all still realistic (Table 4). Interestingly, PEST-estimated parameter values for the three PEST simulations all resulted in almost identical predictions of phytoplankton and zooplankton concentrations as the ad hoc calibrated simulation. PEST appears to be quite robust. We repeated some of the PEST calibrations using different starting values for parameters and with some of the numerical settings of PEST changed and obtained very similar calibrated parameter values. Thus, with proper and careful implementation, PEST offers a viable approach for objective calibration of NEMURO to site-specific monitoring data.

Comparison of the ad hoc and PEST calibrations of NEMURO to the field data (simulation 4) illustrated the potential arbitrariness and sublime decision-making inherent in ad hoc model calibration and a potential pitfall of automatic calibration methods. When presented with a consistent dataset (i.e., model predictions in PEST simulations 1–3), PEST was able to recover known parameter values. When confronted with the field dataset in simulation 4, PEST-estimated and ad hoc calibrated parameter values differed (Table 4) and provided alternative fits or interpretations of the same field data (Fig. 5). These differences in parameter values and fits were important, as they generated somewhat different long-term

dynamics (Fig. 6). Although from a statistical viewpoint the PEST calibration fitted the field data better than the ad hoc calibration (averaged residuals of 0.0081 versus  $-0.11$ ), we consider the ad hoc calibration to be a more useful fit to the field data in this specific situation. Some of the PEST-estimated parameter values were near the extreme of reasonable values, and total phytoplankton were dominated by small phytoplankton. Without explicitly stating it, we realized that during the ad hoc calibration we had decided that large phytoplankton biomass should be similar to or higher than small phytoplankton biomass. This decision was simply based on qualitative information and our experience in applying the NEMURO to other, admittedly not coastal, locations (e.g., Fujii et al., 2002; Yamanaka et al., 2004). We suspect that these sublime decisions occur often with ad hoc calibration, and are one of the major disadvantages to ad hoc approaches. This information on the goal of wanting roughly similar or greater biomass of large phytoplankton was not available to PEST, who only saw field data on total phytoplankton. Thus, PEST quite reasonably opted to fit the other zooplankton groups better than the ad hoc calibration by satisfying the total phytoplankton with higher small phytoplankton and lower large phytoplankton. It is not clear how to avoid these types of decisions in ad hoc calibration or how to incorporate this kind of qualitative information into an automatic calibration method such as PEST. In particular, PEST allows for easy specification of constraints on parameters, but not field data. Boyle et al. (2000) suggest combining manual and automatic calibration approaches by applying automatic calibration separately to subsets of desired model behaviors, and then graphically examining the tradeoffs between improving the fit on some behaviors at the expense of decreased fit for other behaviors.

A variety of additional PEST simulations that used the WCVI dataset beyond the PEST simulation 4 reported in this paper was attempted. These included allowing all 57 parameters to be varied, and providing PEST with small and large phytoplankton concentrations specified separately (rather than as a single variable of total phytoplankton). All of these resulted in similar fits to the field data with one or more parameters at their minimum or maximum allowed values. One interpretation of these results is that the WCVI field dataset, as we constructed it from multiple sources and years of data collection, was inconsistent with the NEMURO model. PEST can estimate parameter values for small and large phytoplankton using total phytoplankton if provided with consistent data on total phytoplankton. Yet, PEST was unable to satisfactorily calibrate to the field data. This result suggests that we should revisit how we constructed the idealized time series dataset.

Calibration results and estimated parameter values can vary greatly if aggregated calibration variables (e.g., total phytoplankton) are used, and the benefit of automatic calibration methods being objective also makes formal inclusion of some constraints difficult. We could have continued to make the PEST calibration more complicated by weighting field monitored variables differently and down-weighting the troughs of small zooplankton, but eventually we would be approaching another version of an ad hoc calibration. Our results illustrate the importance of understanding the quality of the calibration dataset. Robust parameter estimation, using either ad hoc or automatic calibration methods, depends on the quality and

consistency of the calibration dataset and its compatibility with the structure of the model. By comparing PEST and ad hoc calibrations we clearly demonstrated that calibration of NEMURO is best achieved with either explicitly defined time series data on each of the phytoplankton and zooplankton state variables; if one must use total (summed) state variables then it is critical to understand the quality of the aggregated time series and to formally describe the behaviors desired from the model.

There are many features of PEST that were not utilized in this paper. As a consequence of the many model runs performed as part of the parameter estimation, PEST accumulates information on parameter sensitivities (Jacobian matrix) and covariance (correlation) among parameters. For example, PEST calibration 3 that varied 57 parameters and use the ad hoc calibration as observed data showed that the two maximum photosynthesis rates ( $V_{maxs}$ ,  $V_{maxl}$ ) and the maximum grazing rate of large zooplankton on large phytoplankton ( $G_{maxlpl}$ ) were most important in affecting model predictions. The relative sensitivities, which measure the composite change in model predictions due to a fractional change in the parameter (Doherty, 2004), exceeded 1.0 for these three parameters. Although performed in a different context, Kuroda and Kishi (2004) and Yoshie et al. (2007) also generally concluded that maximum photosynthesis and maximum grazing rates were important parameters. Because of the large number of adjustable parameters relative to the amount of calibration data, there were many correlations among parameters. The PEST-estimation parameter correlation matrix showed about half the pairs of parameters had correlation coefficients  $>0.8$ , which was consistent with a matrix condition number of 723,964 ( $>10,000$  signifies a matrix approaching singularity due to parameter insensitivities or correlations, Doherty, 2004). Other features of PEST not utilized in this paper include differential weighting of state variables in the objective function, accommodation of prior information to constrain parameter values, and prediction mode that allows for quantification of uncertainty due to non-uniqueness of parameter solutions. Many of these features are also common to other search-based parameter estimation methods (e.g., Matear, 1995).

Kuroda and Kishi (2004) also applied an objective calibration method to a one-box version of NEMURO using monitoring data for Station A-7. They allowed eight parameters determined from a previous sensitivity analysis to be important to be varied, and used the conjugate gradient method to minimize the squared difference between predicted and observed values. Kuroda and Kishi (2004) concluded that while the objectively estimated parameter values generate model predictions closer to the observed data than a set of first-guess parameter values, several discrepancies between predicted and observed values remained. We view our analysis here as a continuation in the same spirit of the data assimilation analysis of Kuroda and Kishi (2004).

We arrived at similar conclusions as Fennel et al. (2001) that sensitivity analysis and the quality of the assembled field data were critical to effective parameter estimation. Fennel et al. (2001) applied a similar automatic parameter estimation approach as PEST to a simple nitrogen-phytoplankton-zooplankton model. They con-

cluded that sensitivity analysis was a critical accompaniment to parameter estimation for revealing inadequacies in the formulation of the optimization problem, and that their model was inconsistent with the available field data. Additional sensitivity analysis of NEMURO (e.g., Yoshie et al., 2007), or at least a systematic approach to identify which parameters to vary, would help constrain PEST in future applications.

We recommend that automatic calibration methods, such as PEST, be used for application of the NEMURO model to new locations. Ideally, field data will allow for specification of time series for each phytoplankton and zooplankton state variable, and the available field data are sufficient to be kept separate by their year of collection. In this situation, PEST will provide an objective, defensible, and repeatable way to calibrate the many parameters of the NEMURO model. Unfortunately, phytoplankton biomass is often reported as chlorophyll, rather than broken down into functional groups or size-classes. If the available data are insufficient for specification of each state variable, then ad hoc calibration will likely be needed to allow for inclusion of qualitative decisions about model fit (or mis-fit). We recommend applying PEST in this situation as well. Calibration using PEST provides a method for better understanding the causes for data-model mis-matches, and for providing an alternative calibration to ad hoc in order to clearly understand the decisions underlying the ad hoc calibration.

We are using PEST as part of our broader application of NEMURO coupled with herring and saury bioenergetics model to compare geographical differences in fish growth and its link to climate (Megrey et al., this issue; Rose et al., this issue). We use PEST to systematically calibrate NEMURO to each location, and compare predicted herring and saury size-at-age to assess the importance of environmental differences in explaining patterns of fish growth among locations. We have learned that the data synthesis and time series preparation is critical to obtaining a robust calibration of NEMURO. Ultimately, we hope to use the coupled models to examine the effects of regime shifts and global climate change on herring growth.

---

## Acknowledgements

We would like to thank the APN (Asia Pacific Network), North Pacific Marine Science Organization (PICES), the Global Ocean Ecosystem Dynamics Program (GLOBEC), the Heiwa-Nakajima Foundation of Japan, the Japan International Science and Technology Exchange Center, the City of Nemuro (Japan), and the Fisheries Research Agency (FRA) of Japan for sponsoring a series of workshops that resulted in the development of the NEMURO model and its coupling to fish growth dynamics. BAM's participation in this research is noted as contribution FOCI-0536 to NOAA's Fisheries-Oceanography Coordinated Investigations.

## REFERENCES

- 
- Aita, M.N., Yamanaka, Y., Kishi, M.J., 2003. Effects of ontogenetic vertical migration of zooplankton on annual primary



- production—using NEMURO embedded in a general circulation model. *Fish. Oceanogr.* 12, 284–290.
- Al-Abed, N.A., Whiteley, H.R., 2002. Calibration of the Hydrological Simulation Program Fortran (HSPF) model using automatic calibration and geographical information systems. *Hydrol. Process.* 16, 3169–3188.
- Baginska, B., Milne-Home, W., Cornish, P.S., 2003. Modelling nutrient transport in Currency Creek, NSW with AnnAGNPS and PEST. *Environ. Modell. Softw.* 18, 801–808.
- Boyle, D.P., Gupta, H.V., Sorooshian, S., 2000. Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods. *Water Resour. Res.* 36, 3663–3674.
- Doherty, J., 2004. PEST: Model Independent Parameter Estimation User Manual, 5th ed. Watermark Numerical Computing, Brisbane, Australia.
- Doherty, J., Johnston, J.M., 2003. Methodologies for calibration and predictive analysis of a watershed model. *J. Am. Water Resour. Assoc.*, 251–265.
- Eckhardt, K., Arnold, J.G., 2001. Automatic calibration of a distributed catchment model. *J. Hydrol.* 251, 103–109.
- Fennel, K., Losch, M., Schroter, J., Wenzel, M., 2001. Testing a marine ecosystem model: sensitivity analysis and parameter optimization. *J. Mar. Syst.* 28, 45–63.
- Freedman, V.L., Lopes, V.L., Hernandez, M., 1998. Parameter identifiability for catchment-scale erosion modelling: a comparison of optimization algorithms. *J. Hydrol.* 207, 83–97.
- Friedrichs, M.A.M., 2002. Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the central equatorial Pacific Ocean. *Deep Sea Res. II* 49, 289–319.
- Fujii, M., Nojiri, Y., Yamanaka, Y., Kishi, M.J., 2002. A one-dimensional ecosystem model applied to time-series Station KNOT. *Deep Sea Res. II* 49, 5441–5461.
- Ito, S., Kishi, M.J., Kurita, Y., Oozeki, Y., Yamanaka, Y., Megrey, B.A., Werner, F.E., 2004. Initial design for a fish bioenergetics model of Pacific saury coupled with a lower trophic ecosystem model. *Fish. Oceanogr.* 13 (Suppl. 1), 111–124.
- Harmon, R., Challenor, P., 1997. A Markov chain Monte Carlo method for estimation and assimilation into models. *Ecol. Model.* 101, 41–59.
- Kishi, M.J., Kashiwai, M., Ware, D.M., Megrey, B.A., Eslinger, D.L., Werner, F.E., Aita, M.N., Azumaya, T., Fujii, M., Hashimoto, S., Huang, D., Iizumi, H., Ishida, Y., Kang, S., Kantakov, G.A., Kim, H.-C., Komatsu, K., Navrotsky, V.V., Smith, S.L., Tadokoro, K., Tsuda, A., Yamamura, O., Yamanaka, Y., Yokouchi, K., Yoshie, N., Zhang, J., Zuenko, Y.I., Zvalinsky, V.I., 2007. NEMURO—Introduction to a lower trophic level model for the North Pacific marine ecosystem. *Ecol. Modell.* 202, 12–25.
- Kishi, M.J., Okunishi, T., Yamanaka, Y., 2004. A comparison of simulated particle fluxes using NEMURO and other ecosystem models in the western North Pacific. *J. Oceanogr.* 60, 63–73.
- Kuroda, H., Kishi, M.J., 2004. A data assimilation technique applied to “NEMURO” for estimating parameter values. *Ecol. Model.* 172, 69–85.
- Mackas, D.L., 1992. Seasonal cycle of zooplankton off southwestern British Columbia: 1979–1989. *Can. J. Fish. Aquat. Sci.* 49, 903–921.
- Mackas, D.L., Peterson, W.T., Zamon, J.E., 2004. Comparisons of interannual biomass anomalies of zooplankton communities along the continental margins of British Columbia and Oregon. *Deep Sea Res. II* 51, 875–896.
- Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J. Hydrol.* 235, 276–288.
- Matear, R.J., 1995. Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at Station P. *J. Mar. Res.* 53, 571–607.
- Megrey, B.A., Rose, K.A., Klumb, R.A., Hay, D.E., Werner, F.E., Eslinger, D.L., Smith, S.L., this issue. A bioenergetics population dynamics model of Pacific herring (*Clupea pallasii*) coupled to NEMURO: dynamics, description, validation and sensitivity analysis. *Ecol. Modell.*
- Mulligan, A.E., Brown, L.C., 1998. Genetic algorithms for calibrating water quality models. *J. Environ. Eng.*, 202–211.
- Robinson, A.R., Lermusiaux, P.F.J., 2002. Data assimilation for modeling and predicting coupled physical–biological interactions in the Sea. In: Robinson, A.R., McCarthy, J.J., Rothschild, B.J. (Eds.), *The Sea*, vol. 12. John Wiley and Sons, New York, pp. 475–536 (Chapter 12).
- Robinson, C.L.K., Ware, D.M., Parsons, T.R., 1993. Simulated annual plankton production in the northeastern Pacific coastal upwelling domain. *J. Plankton Res.* 15, 161–183.
- Robinson, C.L.K., Ware, D.M., 1999. Simulated and observed response of the southwest Vancouver Island pelagic ecosystem to oceanic conditions in the 1990s. *Can. J. Fish. Aquat. Sci.* 56, 2433–2443.
- Rose, K.A., Cook, R.B., Brenkert, A.L., Gardner, R.H., Hettelingh, J.P., 1991. Systematic comparison of ILWAS, MAGIC, and ETD watershed acidification models. 1. Mapping among model inputs and deterministic results. *Water Resour. Res.* 27, 2577–2589.
- Rose, K.A., Werner, F.E., Megrey, B.A., Aita, M.N., Yamanaka, Y., Hay, D.E., Schweigert, J., Foster, M.B., this issue. Simulated herring growth responses in the northeastern Pacific to historic temperature and zooplankton conditions generated by the 3-dimensional NEMURO nutrient–phytoplankton–zooplankton model. *Ecol. Modell.*
- Schartau, M., Oschlies, A., Willebrand, J., 2001. Parameter estimates of a zero-dimensional ecosystem model applying the adjoint method. *Deep Sea Res. II* 48, 1769–1800.
- Smith, S.L., Yamanaka, Y., Kishi, M.J., 2005. Attempting consistent simulations of Stn. ALOHA with a multi-element ecosystem model. *J. Oceanogr.* 61, 1–23.
- Swartzman, G.L., Kaluzny, S.P., 1987. *Ecological Simulation Primer*. Macmillan Publishing Company, New York.
- Thomson, R.E., Fine, I.V., 2003. Estimating mixed layer depth from oceanic profile data. *J. Atmos. Ocean. Technol.* 20, 319–329.
- Vallino, J.J., 2000. Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *J. Mar. Res.* 58, 117–164.
- Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S., 2003. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters. *Water Resour. Res.* (SWC) 39, 1-1-1-15.
- Wainwright, T.C., Feinberg, L.R., Hooff, R.C., Peterson, W.T., this issue. A comparison of two lower trophic models for the California Current system. *Ecol. Modell.*
- Wang, S.-B., 1998. Comparative studies of the life history and production potential of bay anchovy *Anchoa mitchilli* and northern anchovy *Engraulis mordax*: an individual-based modeling approach. PhD Dissertation. University of South Alabama, Mobile.
- Werner, F.E., Ito, S., Megrey, B.A., Kishi, M.J., this issue. Synthesis of the NEMURO model studies and future directions of marine ecosystem modeling. *Ecol. Modell.*
- Yamanaka, Y., Yoshie, N., Fujii, M., Aita-Noguchi, M., Kishi, M.J., 2004. An ecosystem model coupled with nitrogen–silicon–carbon cycles applied to station A7 in the northwest Pacific. *J. Oceanogr.* 60, 227–241.
- Yoshie, N., Yamanaka, Y., Kishi, M.J., Saito, H., 2003. One dimensional ecosystem model simulation of the effects of

- vertical dilution by the winter mixing on the spring diatom bloom. *J. Oceanogr.* 59, 563–571.
- Yoshie, N., Yamanaka, Y., Rose, K.A., Eslinger, D.L., Ware, D.M., Kishi, M.J., 2007. Parameter sensitivity study of the NEMURO lower trophic level marine ecosystem model. *Ecol. Modell.* 202, 26–37.
- Zyvoloski, G., Kwicklis, E., Eddebarh, A.A., Arnold, B., Faunt, C., Robinson, B.A., 2003. The site-scale saturated zone flow model for Yucca Mountain: calibration of different conceptual models and their impact on flow paths. *J. Contam. Hydrol.* 62/63, 731–750.