



Development of a logistic regression model for the prediction of toxigenic *Pseudo-nitzschia* blooms in Monterey Bay, California

Jenny Q. Lane^{1,*}, Peter T. Raimondi², Raphael M. Kudela¹

¹Department of Ocean Sciences, University of California, 1156 High Street, Santa Cruz, California 95064, USA

²Department of Ecology and Evolutionary Biology, University of California, Center for Ocean Health, 100 Shaffer Road, Santa Cruz, California 95064, USA

ABSTRACT: Blooms of the diatom genus *Pseudo-nitzschia* have been recognized as a public health issue in California since 1991 when domoic acid, the neurotoxin produced by toxigenic species of *Pseudo-nitzschia*, was first detected in local shellfish. Although these blooms are recurring and recognized hazards, the factors driving bloom proliferation remain poorly understood. The lack of long-term field studies and/or deficiencies in the scope of environmental data included within them hinders the development of robust forecasting tools. For this study, we successfully developed predictive logistic models of toxigenic *Pseudo-nitzschia* blooms in Monterey Bay, California, from a multi-project dataset representing 8.3 yr of sampling effort. Models were developed for year-round (annual model) or seasonal use (spring and fall–winter models). The consideration of seasonality was significant: chlorophyll *a* (chl *a*) and silicic acid were predictors in all models, but period-specific inclusions of temperature, upwelling index, river discharge, and/or nitrate provided significant model refinement. Predictive power for ‘unknown’ (future) bloom cases was demonstrated at $\geq 75\%$ for all models, out-performing a chl *a* anomaly model, and performing comparably to, or better than, previously described statistical models for *Pseudo-nitzschia* blooms or toxicity. The models presented here are the first to have been developed from long-term (>1.5 yr) monitoring efforts, and the first to have been developed for bloom prediction of toxigenic *Pseudo-nitzschia* species. The descriptive capacity of our models places historical and recent observations into greater ecological context, which could help to resolve historical alternation between the implication of freshwater discharge and upwelling processes in bloom dynamics.

KEY WORDS: *Pseudo-nitzschia* · Predictive model · Logistic regression · Harmful algal bloom · Phytoplankton monitoring · Domoic acid

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Harmful algal blooms (HABs) can have severe deleterious consequences for local industry (e.g. shellfish, tourism), public health, and ecosystem health. In addition, the incidence of HABs appears to be increasing in both frequency and intensity (Hallegraeff 1993, Anderson et al. 2002, Glibert et al. 2005). This trend, and its potential to inflict rising economic and societal costs, has encouraged the development of HAB forecasting tools in recent years (Schofield et al. 1999,

Johnsen & Sakshuag 2000, Fisher et al. 2003). Many of these efforts have focused on the prediction and monitoring of dinoflagellate blooms and associated red tides, and successful prediction models for these types of HABs span a wide range of modeling approaches and complexity. One of the simplest approaches utilizes satellite-derived chlorophyll anomalies to identify potentially harmful blooms (e.g. Allen et al. 2008) or even species-specific blooms (Tomlinson et al. 2004). Issues associated with satellite-derived models (non-specificity, infrequent data) can be overcome by com-

*Email: jquay@ucsc.edu

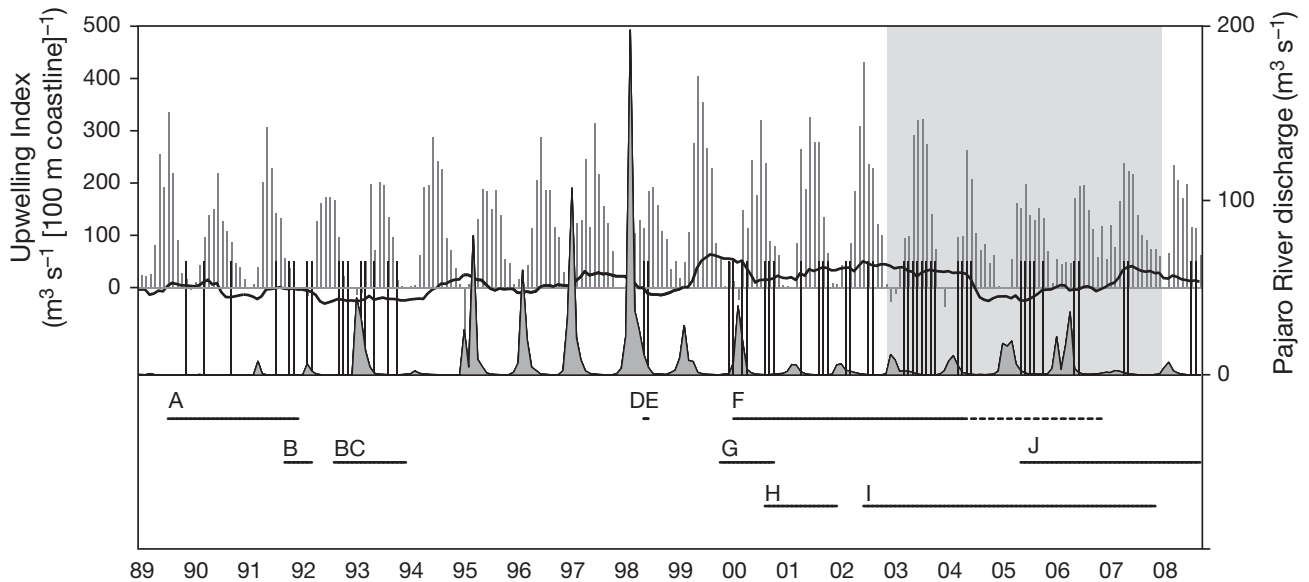


Fig. 1. Time series of bloom events, as reported in published literature and in the modeling dataset used here. The time periods addressed by the various studies are indicated below the plot and are as follows: A: Buck et al. (1992); B: Walz et al. (1994); C: Walz (1995); D: Scholin et al. (2000); E: Trainer et al. (2000); F: Jester et al. (2009); G: Lefebvre et al. (2002b); H: Goldberg (2003); I: Center for Integrated Marine Technologies model dataset (present study); J: California Program for Regional Enhanced Monitoring for PhycoToxins model dataset (present study). Grey shading: time-frame of data inclusion for the models developed in the present study. Vertical black bars: months for which blooms of toxicogenic *Pseudo-nitzschia* (cell concentration $\geq 10\,000$ cells l^{-1}) were reported. A 12 mo moving average of monthly upwelling anomaly (black line), the monthly upwelling index for $36^{\circ}N$, $122^{\circ}W$ (grey bars), and monthly mean Pajaro River discharge (area plot) are shown. Within the time-frame of Jester et al. (2009) (F), note the negative values of the upwelling anomaly and the relatively high monthly mean river discharges over the span of summer 2004 through 2006 (F; dashed line); this period was identified by Jester et al. (2009) as a period of severely decreased *Pseudo-nitzschia* abundance

binning satellite data with other predictors, as is done in the Southwest Florida (USA) operational forecast for *Karenia brevis*. This model includes satellite data, wind predictions, and rule-based modeling to improve forecasting success (Stumpf et al. 2009). A similar approach integrating multiple environmental datasets was used for the European Harmful Algal Bloom Expert System (HABES); this predictive modeling approach uses 'fuzzy logic' to identify blooms of *Nodularia spumigena*, *Dinophysis* spp., *Alexandrium minutum*, *K. mikimotoi*, and *Phaeocystis globosa* (Blauw et al. 2006). The authors state that fuzzy logic bridges the gap between purely empirical (statistical) predictions and fully deterministic models. Finally, seasonal initiation of HAB events and spatial/temporal distribution have been successfully predicted using fully coupled deterministic physical-biological models in the Gulf of Maine (USA) for *A. fundyense* (McGillicuddy et al. 2005). In contrast to these and other efforts, relatively little predictive skill has been developed for HABs of diatom species.

Toxigenic species of the diatom *Pseudo-nitzschia* are producers of domoic acid, which can cause neurotoxic poisoning in humans (Addison & Stewart 1989, Bates et al. 1989), marine mammals (Lefebvre et al. 1999, 2002a, Scholin et al. 2000, Kreuder et al. 2005), and

birds (Fritz et al. 1992, Beltrán et al. 1997). Since initial documentation in 1991, HABs of toxicogenic *Pseudo-nitzschia* have occurred in Monterey Bay, California, with regularity (Buck et al. 1992, Fritz et al. 1992, Work et al. 1993, Scholin et al. 2000, Trainer et al. 2001); the bulk of published *Pseudo-nitzschia* bloom data was generated through episodic, generally stand-alone, research projects undertaken in reaction to these periodic events (Fig. 1). Because these studies were relatively short and episodic in nature, they forcibly relied on circumstantial observations or single-variable correlations for identification of environmental conditions conducive to bloom formation; the constraints associated with this approach led to calls for long-term monitoring approaches (Trainer et al. 2000, Bates & Trainer 2006).

Despite the lack of long-term data, eutrophication via terrestrial freshwater runoff (Bird & Wright 1989, Trainer et al. 1998, Scholin et al. 2000), fluctuations in nutrient ratios (Marchetti et al. 2004), and upwelling processes (Buck et al. 1992, Trainer et al. 2000, Anderson et al. 2006) were implicated as prominent causative factors in historical literature, and our modeling design was developed with this historical ecological perspective in mind. While our data used for model development could not extend over the full time period

represented within the literature due to sampling and methodological inconsistencies, the range of ecological circumstances addressed are not unlike those encountered and implicated previously as triggers for HAB events (Fig. 1).

Our efforts follow 2 previous modeling studies that used shorter duration datasets. The first (Blum et al. 2006) was an attempt to model cellular domoic acid in a toxigenic strain of *Pseudo-nitzschia* (*Pseudo-nitzschia pungens* f. *multiseriis*). In that study, 4 models were developed: 2 linear models demonstrated 'good predictive ability', but were developed from laboratory data that failed to address the scope of nutrient concentrations and ratios encountered in the field and were therefore not appropriate for use with field data. A third linear model and a logistic regression model were developed from combined laboratory data and field data collected from monospecific blooms of *P. pungens* f. *multiseriis* off the coasts of Prince Edward Island (Canada), and Washington State (USA). Split-sample validations of these models (75% data used for model development; 25% reserved for model validation) demonstrated their 'adequate reliability', but the limited amount of field data (N = 46) and the predominance of restrictive laboratory data within the modeling dataset left the applicability of these models undetermined.

The second modeling study (Anderson et al. 2009) developed linear regression (hindcast) models of *Pseudo-nitzschia* blooms, particulate domoic acid, and cellular domoic acid, from: (1) a 'full' (remotely sensed and *in situ*) suite of predictor variables and (2) a 'remote-sensing only' suite of predictor variables. This study was limited in the amount of data available for model development (N = 72 to 89), but provided preliminary insight into *Pseudo-nitzschia* bloom mechanisms, including macronutrient control. Both model sets presented by Anderson et al. (2009) demonstrated high rates of false negative predictions, presumably due to the relatively limited dataset.

Here, we develop logistic regression models of toxigenic *Pseudo-nitzschia* blooms in Monterey Bay, California. This modeling exercise had 3 goals: (1) to develop *Pseudo-nitzschia* bloom models that are straightforward and useful in their application towards bloom monitoring, (2) through model development, to identify environmental variables that are significant factors in bloom incidence, and (3) to test the recurrence of these significant environmental variables in the previous *Pseudo-nitzschia* models described by Anderson et al. (2009) and Blum et al. (2006). The previous modeling efforts and this one are not wholly consistent in terms of scope, evaluated variables, or specific aim: Anderson et al. (2009) developed models of 'generic' *Pseudo-nitzschia* blooms, cellular domoic

acid, and particulate domoic acid from a 1.5 yr dataset collected from the Santa Barbara Channel, while Blum et al. (2006) developed models of particulate domoic acid from a mixture of experimental and field data. These previous studies and our efforts clearly differ in their region of interest and specific model subject. In the context of the present study, these disparities are an advantage, in that they allow for inter-model comparison capable of identifying factors that are likely to be universally significant to *Pseudo-nitzschia* bloom incidences and to the introduction of domoic acid into the marine environment through bloom proliferation. Thus, our model and the comparison of these 3 efforts should help to identify a common set of variables useful for predictive modeling of *Pseudo-nitzschia* in similar systems, such as major eastern boundary current regimes (Kudela et al. 2005).

We present 3 logistic regression models of toxigenic *Pseudo-nitzschia* blooms in Monterey Bay, California, as they occur throughout the year (annual model) and seasonally (spring and fall–winter models). A total of 31 environmental variables were evaluated, and 6 variables were identified as statistically significant for bloom prediction. This work is the first to present robust *Pseudo-nitzschia* bloom models developed from long-term monitoring data, and the first to evaluate eutrophication processes and seasonality in the prediction of *Pseudo-nitzschia* bloom incidences.

MATERIALS AND METHODS

Compilation of the model dataset. We compiled a dataset from publications that included *Pseudo-nitzschia* cell counts for Monterey Bay (Buck et al. 1992, Walz et al. 1994, Walz 1995, Villac 1996, Scholin et al. 2000, Goldberg 2003, Lefebvre et al. 2002b). Additional unpublished datasets were provided by Moss Landing Marine Laboratories (MLML), and internally generated through the Center for Integrated Marine Technologies (CIMT) and through the California Program for Regional Enhanced Monitoring for PhycoToxins (Cal-PreEMPT). Details on sampling and analytical methods for internally generated datasets are provided.

We obtained 2099 discrete cases from the above sources, 1156 of which were from surface waters (depth ≤ 5 m). All of the data were assessed to ensure methodological consistency, specifically: (1) unbiased sample collection and (2) true concurrency in environmental and *Pseudo-nitzschia* sampling. Of the 1071 cases remaining, 576 contained cell counts of toxigenic *Pseudo-nitzschia*. Not all data contained the same suite of environmental variables. For finalization of the modeling dataset, it was necessary to evaluate which data

were sufficiently complete, i.e. evaluate the minimal combination of variables sufficient for the development of a successful model. The receiver operating characteristic (ROC) was used to conduct this evaluation. ROC is a measure of model fit that scales like a traditional (US) academic point system (<0.6 = poor; 0.6 to 0.7 = fair; 0.7 to 0.8 = good; 0.8 to 0.9 = very good; >0.9 = excellent). Models developed from single, single and universally available (i.e. river discharge, upwelling index), and pairs of predictor variables failed to achieve 'very good' model fit accuracy. To achieve this level of accuracy, model development required concurrent macronutrient, chlorophyll *a* (chl *a*), and temperature variables in combination (Tables 1 & 2). Final inclusion of cases for the models presented therefore required sample collection from Monterey Bay surface waters, and toxigenic *Pseudo-nitzschia* cell counts (*P. multiseriata* and/or *P. australis*) with concurrent environmental measurements of seawater temperature, chl *a*, and macronutrients.

Internal data: sample collection. Samples were collected monthly from June 2002 to November 2007 from 11 stations throughout Monterey Bay as part of the CIMT project. PVC Niskin bottles (10 l volume fitted with silicone rubber band strings) mounted on an instrumented rosette were used to collect water from 5 m depth. Surface samples were collected from 2 stations by PVC bucket. Temperature data were obtained from a Seabird SBE-19 CTD deployed concurrently with water sampling.

Table 1. Evaluations of independent variable(s) as predictor variables were performed using all compiled literature and field data of toxic *Pseudo-nitzschia* in Monterey Bay at a depth ≤ 5 m (N = 576). The receiver operating characteristic (ROC) is a measure of model fit accuracy, where <0.6 = poor, 0.6 to 0.7 = fair, 0.7 to 0.8 = good, 0.8 to 0.9 = very good, and >0.9 is considered excellent. Inclusion of macronutrient, seawater temperature, and chlorophyll *a* as predictor variables was necessary to achieve 'very good' model fit accuracy.

A key to variable names is provided in Table 2

Independent variable	N (cases)	Cases omitted	ROC
Salinity	427	149	0.462
Temp	493	83	0.573
ln(silicic acid)	516	60	0.614
ln(chl <i>a</i>)	497	79	0.618
ln(chl <i>a</i>), temp	473	103	0.638
ln(chl <i>a</i>), upwelling	492	84	0.638
ln(silicic acid), upwelling	516	60	0.713
ln(silicic acid), ln(chl <i>a</i>)	444	132	0.757
ln(nitrate), temp, ln(chl <i>a</i>)	419	157	0.766
ln(silicic acid), temp	438	138	0.785
ln(silicic acid), temp, ln(chl <i>a</i>)	422	154	0.848

Table 2. Complete list of the variables evaluated as independent (predictor) variables in the logistic regression models. X: all environmental variables and ratios, excluding temperature

Independent variable	Abbreviation	Units
Seawater temperature	Temp	°C
Total chlorophyll <i>a</i>	Chl <i>a</i>	$\mu\text{g l}^{-1}$
Nitrate	Nitrate	μM
Silicic acid	Silicic acid	μM
Ortho-phosphate	Phosphate	μM
Silicic acid (nitrate) ⁻¹	Silicic acid:nitrate	
Nitrate (silicic acid) ⁻¹	Nitrate:silicic acid	
Ortho-phosphate (nitrate) ⁻¹	Phosphate:nitrate	
Nitrate (ortho-phosphate) ⁻¹	Nitrate:phosphate	
Ortho-phosphate (silicic acid) ⁻¹	Phosphate:silicic acid	
Silicic acid (ortho-phosphate) ⁻¹	Silicic acid:phosphate	
Pajaro River discharge	Pajaro River	$\text{m}^3 \text{s}^{-1}$
San Lorenzo River discharge	San Lorenzo River	$\text{m}^3 \text{s}^{-1}$
Soquel River discharge	Soquel River	$\text{m}^3 \text{s}^{-1}$
Salinas River discharge	Salinas River	$\text{m}^3 \text{s}^{-1}$
Bakun upwelling index	Upwelling	$\text{m}^3 \text{s}^{-1}$
Ln(X + 1)	Ln(X)	

Shore-based surface samples were collected weekly from May 2005 to April 2008 from the Santa Cruz Municipal Wharf (36° 57.48' N, 122° 1.02' W) as part of the Cal-PreEMPT project using a PVC bucket or by integration of water samples collected from 3 discrete depths (0, 1.5, and 3 m) with a FieldMaster 1.75 l basic water bottle. Temperature was measured in the field by digital thermometer immediately following sample retrieval.

River discharge rates for the Salinas, San Lorenzo, Soquel, and Pajaro Rivers were obtained from the United States Geological Survey National Water Information System (<http://waterdata.usgs.gov/nwis/>). Bakun daily upwelling index values for the Monterey Bay region (36°N, 122°W) were obtained from the National Oceanographic and Atmospheric Administration Pacific Environmental Research Division (www.pfeg.noaa.gov/products/PFEL/).

Internal data: analytical methods. Samples for chl *a* were collected in duplicate and filtered onto uncombusted glass-fiber filters (Whatman GF/F) and processed using the non-acidification method (Welschmeyer 1994). Macronutrients (Nitrate plus nitrite [hereafter referred to as nitrate], silicic acid and ortho-phosphate) were stored frozen prior to analysis with a Lachat Quick Chem 8000 Flow Injection Analysis system using standard colorimetric techniques (Knepel & Bogren 2001, Smith & Bogren 2001a,b). *Pseudo-nitzschia* species identification and enumeration utilized species-specific large subunit rRNA-targeted

probes following standard protocols (Miller & Scholin 1998). Samples were enumerated with a Zeiss Standard 18 compound microscope equipped with a fluorescence Illuminator 100 (Zeiss). Duplicate filters were prepared for each species, and the entire surface area of each filter was considered in counting.

Model development. Logistic regression models were developed using MYSTAT Version 12.02.11. Logistic modeling is appropriate when the dependent variable is dichotomous (e.g. 0/1). Since our dataset contained continuous data of *Pseudo-nitzschia* abundance, logistic modeling required concatenation of *Pseudo-nitzschia* abundance data into a new dichotomous dependent variable (bloom_nonbloom), using a defined bloom threshold of 10 000 toxigenic *Pseudo-nitzschia* cells l^{-1} (Lefebvre et al. 2002b, Fehling et al. 2006, Howard et al. 2007, Jester et al. 2009). Similar model results were obtained (not shown) when a criteria of 5000 cells l^{-1} was used.

Independent variables evaluated during model development are provided in Table 2. We used an automatic stepwise approach (forward, backward, and bidirectional) to identify the most significant subset. Variable selections were refined to: (1) maximize the rate at which blooms were successfully predicted, (2) minimize the rate of false negative predictions, and (3) maximize model fit accuracy (ROC), while controlling for covariance among the independent variables. Variables exhibiting severe covariation, as determined by variance inflation factors and condition indices, were considered mutually exclusive. Only significant variables ($p < 0.05$) were included in the final models. For the development of the 2 seasonal models, the data were partitioned according to the seasonal periods previously described for Monterey Bay (Pennington & Chavez 2000), while the entire dataset was used for development of the annual model. The final 3 models

are as follows: (1) year-round (annual model), (2) February 14 to June 30 (spring model), and (3) July 1 to February 13 (fall–winter model).

We compare our models to a simple bloom prediction method using chl *a* anomalies. Data from June 2004 to July 2008 were obtained from the LOBOVIZ website (www.mbari.org/lobo/loboviz.htm) for a nearshore mooring in Monterey Bay (M0), and a 30 d median chl *a* anomaly was calculated according to methods previously described for 60 d mean anomalies (Tomlinson et al. 2004, Wynne et al. 2006). The LOBOVIZ website was selected as a data source due to its ease of access and applicability. A median was employed in lieu of a mean, since it has recently been recognized as the generally more appropriate value (R. Stumpf pers. comm.).

Translating probability into prediction: prediction-point assignment. When the model equation is solved, the user is presented with the probability of a bloom occurrence. The degree of probability that can be tolerated is referred to here as the ‘prediction-point’. Where the model solution, bloom probability, is greater than the prediction-point the model predicts a bloom. Conversely, where the model solution is lower than the prediction-point, the probability of a bloom is considered sufficiently low to warrant a non-bloom prediction. The prediction-point must be pre-defined by either: (1) the model developer, for optimization of predictive power, or (2) the model user, for selective risk management. We provide optimized prediction-points for each model and offer guidance for their adjustment. Optimized prediction-points were determined by generating model prediction failure rates over the full range of potential prediction-point assignments (0.000 to 1.000) at 0.005 increments (Fig. 2). The overall failure rate is minimized when the failure to predict blooms and failure to predict non-blooms are simulta-

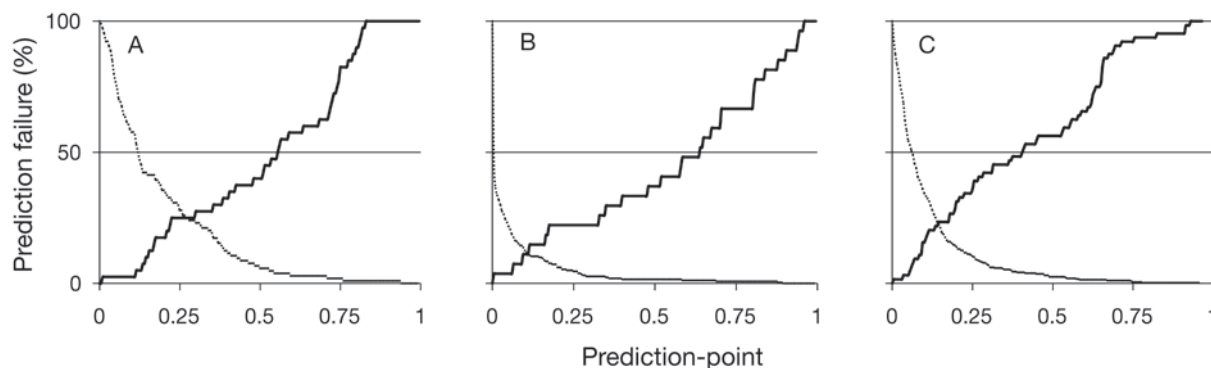


Fig. 2. Prediction failure rates for blooms (solid lines) and non-blooms (broken lines) for the spring (A), fall–winter (B), and annual model (C) along the range of possible prediction-points. The prediction failure rate is defined as the rate at which the model fails to predict a case type (bloom or non-bloom). The overall prediction failure rate is minimized at the optimized prediction-point, where the 2 lines cross

neously minimized. The optimized prediction-point values, therefore, occur where the failure rate curves intersect (Fig. 2).

Model validation. A jackknife cross-validation module was supplied by SYSTAT and used to validate model performance with respect to unknown (future) cases. This method is similar to the split-sample bootstrap validation approach taken by Blum et al. (2006) except that it does not reduce the dataset that can be used for initial model development, and it is an iterative process that allows for N instances of cross-validation against 'unknown' single cases. The cross-validation was run with the model-optimized prediction-points and with 'user-adjusted' prediction-points set according to the historical probability of blooms for Monterey Bay (2002 to 2005), calculated from an independent California Department of Public Health (CDPH) *Pseudo-nitzschia* bloom monitoring dataset. This historical probability is referred to as 'priors'.

RESULTS

After removing those cases from the original (N = 2099) dataset that did not fulfill the specified quality criteria, 506 cases from 2002 to 2008 remained, 74 of which were classified as bloom cases. There was clear seasonality in these data: the rate of bloom incidence was 28% during the spring model period, compared to a rate of 9% for the remainder of the year.

Logistic regression models are of the form:

$$\text{LOGIT}(p) = \ln[p/(1-p)] = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_k z_k \quad (1)$$

where p is the probability of the condition being modeled; here, p represents the probability of a toxigenic *Pseudo-nitzschia* bloom. β_0 is a constant, and $\beta_1, \beta_2, \dots, \beta_k$ are the regression coefficients of z_1, z_2, \dots, z_k , respectively. The year-round (annual) and seasonal (spring and fall-winter) models are as follows:

Annual model

$$\text{LOGIT}(p) = 9.763 - 1.700[\ln(\text{silicic acid})] + 1.132[\ln(\text{chl } a)] - 0.800(\text{temp}) + 0.006(\text{upwelling}) \quad (2)$$

Spring model

$$\text{LOGIT}(p) = 5.835 + 1.398[\ln(\text{chl } a)] - 1.135[\ln(\text{silicic acid})] - 0.549(\text{temp}) \quad (3)$$

Fall-Winter model

$$\text{LOGIT}(p) = 10.832 - 5.026[\ln(\text{Pajaro River})] - 3.893[\ln(\text{silicic acid})] + 1.972[\ln(\text{chl } a)] + 0.652(\text{nitrate}) \quad (4)$$

The regression curve for the spring model is presented for visualization of how the model solution [LOGIT (p)] translates into a bloom probability (p) and, through the

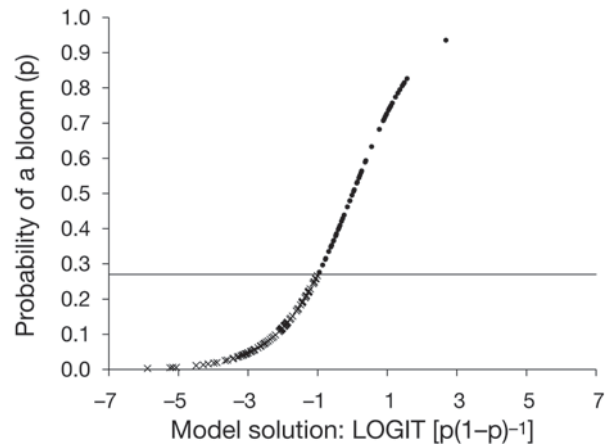


Fig. 3. The spring model logistic regression curve. Cases lying above the optimized prediction-point (probability = 0.275) are predicted as blooms (●); cases lying below the optimized prediction-point are predicted as non-blooms (x)

implementation of a prediction-point, into a bloom or non-bloom prediction (Fig. 3).

The models demonstrated 'very good' to 'excellent' model fit accuracy (Table 3). Other demonstrations of model proficiency include the determination of a model p-value through calculation and evaluation of a likelihood-ratio statistic, and the determination of adjusted R^2 statistics, McFadden's ρ^2 and Nagelkerke's R^2 , for which values >0.2 are indicative of very good model fit (Hensher & Johnson 1981) (Table 3). Each of the models achieved a high level of statistical proficiency with 4 or fewer predictive variables.

Two predictive variables, $\ln(\text{chl } a)$ and $\ln(\text{silicic acid})$, were included in all models. The set of predictive variables used in the annual model and spring model were most similar, differing only in the inclusion of upwelling in the annual model. The fall-winter model is the most disparate of the 3 models with the variable set $\ln(\text{Pajaro River})$, nitrate, $\ln(\text{silicic acid})$ and $\ln(\text{chl } a)$. Of particular note is the omission of temp and upwelling from the fall-winter model, and the inclusion of a river discharge variable, $\ln(\text{Pajaro River})$, and nitrate. All variables included in the models demonstrated extreme to maximum statistical significance.

The annual model did not emerge as an explicit sum of the 2 seasonal models. An annual model with $\ln(\text{Pajaro River})$ and nitrate included (not shown) achieves results very similar to those of the presented annual model, but with slightly improved bloom and non-bloom prediction. These variables, however, were non-significant additions and caused inflation in the standard errors of the other (significant) variable coefficients; this is a general risk assumed when non-significant predictors are included in any model (Menard 1995). The inclusion of additional variables in the

annual model also resulted in unacceptable levels of covariance (condition indices > 30). Dueling complexity and covariance restrictions likely promoted the usefulness of a 'composite' variable, such as upwelling, as a predictor within this, the most temporally comprehensive of the models.

Analysis of model performance at the default prediction-point (0.500) is useful, because it allows for an even comparison of predictive success under equalizing but unrealistic assumptions that: (1) blooms are evenly distributed throughout the year and (2) blooms are expected to occur with as much frequency as non-bloom conditions. As shown in Table 4, the development of seasonal models significantly enhanced pre-

dictive ability: the rate at which blooms were successfully predicted was 16 % (spring model) and 19 % (fall-winter model) greater than for the annual model. The rates of false positive prediction were slightly improved in the seasonal models. The rates of false negative prediction were more unequal among the models, and ranged most significantly between the two seasonal models.

Model performance at optimized prediction-points is summarized in Table 4; the fall-winter model demonstrated the highest rates of case prediction, followed by the annual and spring models, respectively. As with the default prediction-point, the rates of false negative prediction are most disparate among the

seasonal models. The rates of false positive prediction are lowest for the seasonal models, but are increased overall with the implementation of the optimized prediction-points. The relatively high rates of false positive prediction result from a relatively low frequency of non-bloom predictions, which is an artifact of prediction-point optimization.

All of the models were assessed for predictive performance with unknown (future) cases by jackknife validation (Table 5). At the model-optimized prediction-points (Table 5), the rates at which blooms are successfully predicted are more comparable between the spring and annual models, and highest in the fall-winter model. Each model significantly out-performed a null model, improving bloom prediction by as much as 80%. This advantage does not extend to the prediction of non-bloom cases. The discrepancy in bloom versus non-bloom predictive improvement is a result of the model development, which focused on prediction of blooms. The Pearson's chi-squared test statistic for each model indicates extreme significance in the association of modeled predictions and the true outcome of future cases.

The models were also assessed by jackknife cross-validation under conditions simulating the application of 'user-adjusted' prediction-points. In Table 5, the application of 'low' prediction-points set by historical priors provides a demonstration of model performance in a period when future blooms occur with unexpectedly high frequency. The apparently conservative response is in part an artifact of logistic regression: logistic models generally guard against the misclassification of cases

Table 3. Model specifications and diagnostics for the logistic regression models presented in this study. The likelihood-ratio test is a test of the null hypothesis that the predictor variable coefficients are zero (i.e. have no predictive value), and can be evaluated for significance as a deviate chi-squared. McFadden's ρ^2 is a transformation of the likelihood-ratio statistic to mimic an R^2 statistic; values between 0.20 and 0.40 are considered very satisfactory (Hensher & Johnson 1981). The Nagelkerke's R^2 is based on both log likelihood and sample size. ROC: receiver operating characteristic

	Spring	Fall-Winter	Annual
Predictor variables	ln(silicic acid) ln(chl a) Temp	ln(silicic acid) ln(chl a) ln(Pajaro River) Nitrate	ln(silicic acid) ln(chl a) Temp Upwelling
N (total cases)	144	289	422
N (bloom cases)	40	27	64
ROC	0.848	0.943	0.860
Likelihood-ratio statistic	45.885	96.859	102.377
p-value	0.000	0.000	0.000
McFadden's ρ^2	0.270	0.540	0.285
Nagelkerke's R^2	0.394	0.616	0.376

Table 4. Prediction success and failure rates (%) at the default prediction-point of 0.500 and at model-specific optimized prediction-points. A modeled bloom probability higher than the prediction-point results in a bloom prediction. 'False negative' is the rate at which non-bloom predictions were incorrect. 'False positive' is the rate at which bloom conditions were predicted where none existed

	Spring	Fall-Winter	Annual
Default prediction-point			
Prediction-point	0.500	0.500	0.500
Blooms successfully predicted	60	63	44
Non-blooms successfully predicted	94	99	98
False negative	14	4	9
False positive	20	19	24
Model-optimized prediction-point			
Prediction-point	0.275	0.110	0.145
Blooms successfully predicted	75	89	77
Non-blooms successfully predicted	75	89	78
False negative	11	1	5
False positive	46	55	62

Table 5. Jackknife validation results for the logistic regression models at optimized prediction-points, where overall prediction error is minimized and at prediction-points equal to the priors of an independent California Department of Public Health (CDPH) bloom monitoring dataset. Improvement in bloom prediction is relative to the performance of a null model. Square brackets: negative scores

	Spring	Fall–Winter	Annual
Optimized prediction-points			
Prediction-point	0.275	0.110	0.145
Blooms successfully predicted (%)	75	89	77
Non-blooms successfully predicted (%)	76	89	77
False negative (%)	11	1	5
False positive (%)	45	55	62
Improvement in bloom prediction (%)	47	80	62
Improvement in non-bloom prediction (%)	4	[2]	[8]
Pearson's chi-squared (χ^2)	31.78	98.98	74.10
χ^2 p-value	0.000	0.000	0.000
CDPH priors prediction-points			
Prediction-point (priors)	0.101	0.066	0.081
Blooms successfully predicted (%)	98	93	91
Non-blooms successfully predicted (%)	42	82	60
False negative (%)	2	1	3
False positive (%)	61	65	71
Improvement in bloom prediction (%)	70	84	76
Improvement in non-bloom prediction (%)	[30]	[9]	[25]
Pearson's chi-squared (χ^2)	21.31	74.37	54.63
χ^2 p-value	0.000	0.000	0.000

belonging to the under-represented case group, a quality that makes them especially attractive for application in high-risk predictive contexts such as environmental regulation and clinical health (Fan & Wang 1998). The use of the CDPH priors is therefore an appropriate but conservative approach, increasing the probability of correctly identifying blooms ($\geq 91\%$ for all models), while reducing rates of false negative prediction ($\geq 3\%$ for all models), but at a cost to non-bloom prediction.

DISCUSSION

Model application: prediction-point adjustment

We sought to develop and deliver robust predictive models of toxigenic *Pseudo-nitzschia* blooms that were straightforward in their application. Further, we hoped to lay a framework for future modeling studies and independent model application, since this is the first time logistic regression has been applied to *Pseudo-nitzschia* bloom prediction. Providing these models with predetermined, optimized prediction-points satisfies the former; application of the models without end-user adjustment provides a statistically robust method for bloom prediction. Optioning how, and when, the optimized prediction-points can be adjusted satisfies

the latter; the ability to modify the prediction-points grants an opportunity to consider and integrate local bloom ecology, specifically frequency, within the model design.

Statistical models should be developed and implemented while remaining mindful of the system under investigation. In particular, the model should be developed and implemented with consideration of: (1) the general frequency at which blooms occur (the priors), (2) the prediction error rates that are inherent to the model, and (3) the cost of prediction error to the model user. The first of these is taken into account by the designation of an optimized prediction-point. Consideration of Points 2 and 3 is left to the discretion of the model user, since it is only necessary when the risk of a *specific type* of predictive error, rather than *overall* predictive error, needs to be reduced.

Use of a shifted prediction-point, rather than the default prediction-point, should be implemented whenever the probabilities of the 2 outcomes are significantly unequal (Neter et al. 1989). At present, blooms of *Pseudo-nitzschia* are relatively rare occurrences (Fig. 1). Because the probability of a bloom is generally not near 50%, a default prediction-point of 0.500 cannot provide optimized predictive capability. By the same reasoning, if a system generally demonstrates priors that are overwhelmingly different from our assumptions, the prediction-point can be reduced (inflated) to account for the more infrequent (frequent) occurrence, and therefore likelihood, of blooms. Similarly, if the cost of a certain type of incorrect prediction (false positive or false negative) is disproportionately high, the prediction-point can be adjusted to protect from that exaggerated cost. Fig. 2 may be used as a guide for balancing the probability of these errors and controlling their relative costs.

Shifting a prediction-point affects the predictive behavior of the model, always forcing compromise: Fig. 4 illustrates the trade-off between minimizing the number of blooms that the model fails to predict and minimizing the number of non-blooms that are identified as blooms. Reducing the prediction-point minimizes overall failure to predict a bloom by relaxing the criteria for bloom prediction. Conversely, increasing the prediction-point means that the criteria for bloom prediction are more strenuous, and blooms will be forecasted only when they are extraordinarily likely to occur.

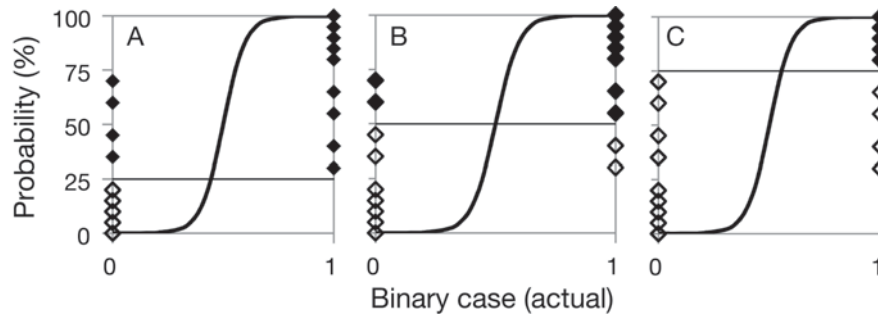


Fig. 4. A schematic diagram of a logistic regression with prediction-points of 0.250 (A), 0.500 (B), or 0.750 (C). Cases lying on the '0' vertical axis are actual non-bloom cases; cases lying on the '1' vertical axis are actual bloom cases. Filled symbols are predicted as blooms; open symbols are predicted as non-blooms. Reducing the prediction-point increases the number of cases that are predicted as blooms overall, maximizing the percent of actual blooms that are successfully predicted but reducing the percent of actual non-blooms that are successfully predicted (A). Increasing the prediction-point has the opposite effect (C)

Comparative model performance

In Table 6, we present the predictive performance of our annual and seasonal models compared to those of: (1) a chl *a* anomaly, (2) linear hindcasting models developed for *Pseudo-nitzschia* blooms in the Santa Barbara Channel (Anderson et al. 2009), and (3) logistic regression models developed for pDA from a combination of field and experimental data (Blum et al. 2006).

Our logistic regression models were developed from the largest dataset to date, and demonstrate a relatively high level of predictive capacity. Our models out-perform the chl *a* anomaly model throughout the year and on a seasonal basis, although the predictive capacity of the chl *a* anomaly model was surprisingly comparable during the spring model period. Interestingly, the chl *a* anomaly model completely failed to predict blooms during the fall–winter model period. We suggest that dinoflagellate blooms, particularly common in Monterey Bay in the fall–winter model period, mask blooms of *Pseudo-nitzschia* otherwise identified by the chl *a* anomaly. Conversely, the chl *a* anomaly works well in the spring model period, when *Pseudo-nitzschia* is more likely to be the dominant bloom organism. It should be noted that the chl *a* anomaly model is advantageous in that it is generally applicable to all potential HABs, particularly 'red tides' (e.g. Kudela et al. 2008b, Ryan et al. 2008), and may therefore be a better model choice when not applied specifically for *Pseudo-nitzschia* bloom prediction.

Table 6. Performance comparisons (%) among the annual and seasonal models developed in the present study and in previous *Pseudo-nitzschia* modeling publications. 'Sensitivity' is the rate at which the binary value '1' cases (blooms or high toxicity) were successfully predicted. 'Specificity' is the rate at which the binary value '0' cases (non-blooms or low toxicity) were successfully predicted. Improvement in bloom prediction is relative to the performance of a null model. Square brackets: negative scores

Dependent variable	Toxigenic <i>Pseudo-nitzschia</i> bloom ^a	<i>Pseudo-nitzschia</i> toxicity ^b	Generic <i>Pseudo-nitzschia</i> bloom ^c	Toxigenic <i>Pseudo-nitzschia</i> bloom ^d
Annual				
Sensitivity	77	77	75	39
Specificity	78	75	93	72
False negative	5	–	25	5
False positive	62	–	7	88
Improvement in bloom prediction	61	–	–	31
N	422	139	75	182
Spring				
Sensitivity	75	–	–	50
Specificity	75	–	–	62
False negative	11	–	–	10
False positive	46	–	–	71
Improvement in bloom prediction	47	–	–	35
N	144	–	–	65
Fall–Winter				
Sensitivity	89	–	–	0
Specificity	89	–	–	77
False negative	1	–	–	3
False positive	55	–	–	100
Improvement in bloom prediction	80	–	–	(2.6)
N	289	–	–	117
^a Present study				
^b Blum et al. (2006)				
^c Anderson et al. (2009)				
^d Study of chl <i>a</i> anomaly (present study)				

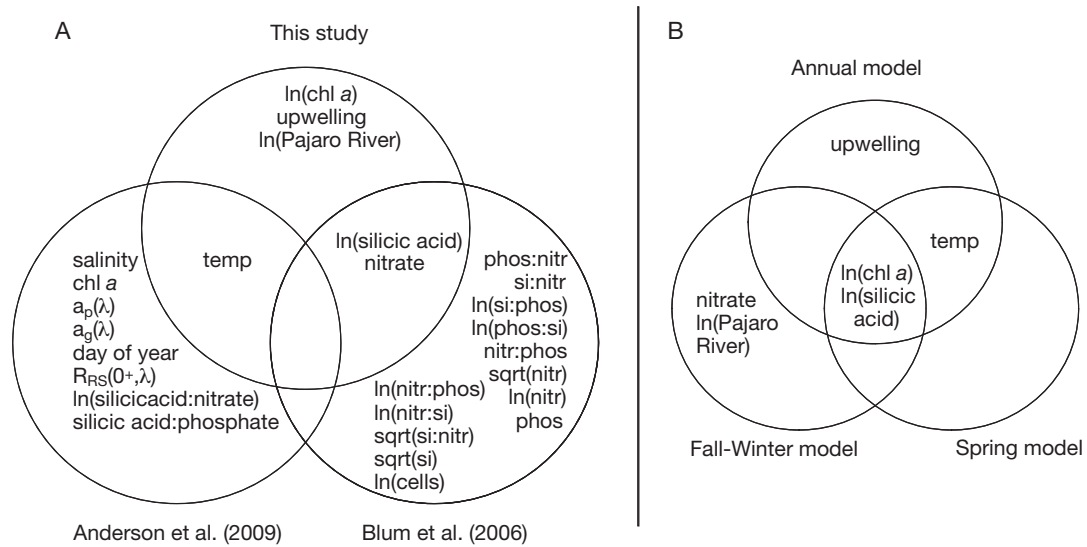


Fig. 5. Venn diagrams illustrating the shared and unique variables included in models of *Pseudo-nitzschia* ecophysiology shown within the present study, Anderson et al. (2009), and Blum et al. (2006) (A) and included in the annual, spring, and fall–winter models (B). Nutrient abbreviations for Blum et al. (2006) are as follows—phos: phosphate; si: silicic acid; nitr: nitrate. Variable abbreviations used in Anderson et al. (2009) are as follows— $a_p(\lambda)$: particulate absorption [412–665 nm]; $a_g(\lambda)$: CDOM absorption [412–665 nm]; $RRS(0^+, \lambda)$: remote-sensing reflectance [412–665 nm]. ‘Sqrt’: square-root operation

Inter-study patterning: recurrent predictor variables

Fig. 5A illustrates the predictor variables that are shared and not shared between the models developed by Blum et al. (2006), Anderson et al. (in press), and in the present study. While the regions of interest and, in some cases, the dependent variable differ between these studies, the similarities and differences shared between the models can provide insight into universal patterns of *Pseudo-nitzschia* ecophysiology and, in turn, indicate which variables may be fundamental to future monitoring and modeling.

Seawater temperature was identified as a significant predictor whenever it was included in a study for evaluation. In all cases, a negative relationship was demonstrated between temperature and the dependent variable. Cold surface temperatures are often associated with upwelling, one of the processes previously identified as a causative factor of *Pseudo-nitzschia* blooms. The direct assessment of the upwelling index was unique to the present study; where it emerged as a predictor variable, it had a weak positive association with *Pseudo-nitzschia* bloom incidence.

Silicic acid (ln-transformed) and nitrate both emerged as predictors in models developed for *Pseudo-nitzschia* toxicity (Blum et al. 2006) and in the models developed here. In both studies, the patterns agree: association with the dependent variable is negative for silicic acid and positive for nitrate. While neither variable emerged as an individual predictor in the models developed by Anderson et al. (2009), a negative relationship was demonstrated between the silicic acid to

nitrate ratio and blooms of *Pseudo-nitzschia*, indicating a possibly confounded negative and positive relationship between blooms and silicic acid and nitrate, respectively.

Additional recurrent patterns are suggested by variables that are related, but not explicitly shared, between the studies. Anderson et al. (2009) chose not to evaluate river discharge as a model variable, but presented a *Pseudo-nitzschia* bloom model and a cellular toxicity model that included particle absorption and absorption of chromophoric dissolved organic matter (CDOM), variables which are associated with significant recent river discharge events (Warrick et al. 2004, 2007). In both models, high particulate absorption was negatively associated with the dependent variable, suggesting a direct negative relationship between high river discharge and *Pseudo-nitzschia* blooms. Our fall–winter model, which addresses the time period in which ‘first flush’ and high discharge events generally occur, also demonstrates a direct negative relationship between river discharge and bloom incidence. The consideration of seasonality when modeling river discharge and blooms and the patterning of blooms and high discharge events through time reveal complexity in this relationship, as discussed in the next subsection.

Intra-study (seasonal) patterning: ecological context and implications

The predictor variables shared and not shared between the annual, spring, and fall–winter models

are presented in Fig. 5B, effectively 'zooming in' on the modeled relationships with a lens of added dimension and ecological context. The 2 most similar models are the annual and spring models; this is not entirely surprising, given that the majority of *Pseudo-nitzschia* blooms occur in the springtime (Fig. 1). Upwelling is the only predictor unique to the annual model; its omission from the spring model may arise from a general predominance of upwelling throughout the spring model period. The independent variables in the spring model exhibited particular propensity for covariation; this would further suggest that *Pseudo-nitzschia* bloom dynamics in Monterey Bay are largely dominated by a specific environmental forcing, i.e. upwelling, over the spring model period.

The fall–winter model includes oceanic periods that are not (by definition) generally dominated by upwelling processes. All of the models, including the fall–winter model, demonstrate that conditions of low silicic acid and concurrently high chl *a* are associated with blooms of toxigenic *Pseudo-nitzschia* in Monterey Bay. The fall–winter model, however, includes 2 unique predictor variables: nitrate (positive coefficient) and Pajaro River discharge (negative coefficient). The inclusion of nitrate in the fall–winter model suggests that the macronutrient control observed by Anderson et al. (2009), specifically the negative relationship between *Pseudo-nitzschia* blooms and the ratio of silicic acid to nitrate, may have been underscored by confounding seasonal relationships. Our results are therefore similar to those presented by Anderson et al. (2009), but are either more specific, due to the explicit assessment of seasonality, or representative of a similar relationship more heavily impacted by eutrophication. Annual dissolved inorganic nitrate loading via terrestrial storm runoff is relatively low in the region addressed by Anderson et al. (2009); however, nitrate input via storm runoff can be significant during winter runoff events (McPhee-Shaw et al. 2007).

The association of fall–winter blooms with conditions of high nitrate suggests that a nitrate eutrophication process is uniquely significant during this period. Notably, the Pajaro River in Monterey Bay introduces disproportionately high nitrate loads (CCLEAN 2006, 2007) on a strictly seasonal basis. In our dataset, blooms within the fall–winter model period occurred during periods of minimal freshwater discharge, while blooms within the spring model period occurred during periods of decreasing river discharge following a 'flush' event (Fig. 1). We also observed this pattern within the broader time-series, in which blooms are generally not associated with peak discharge events and occur either with the declining shoulder of a high river discharge event or within a period marked by minimal discharge (Fig. 1; note that a 4 yr period of

relatively high discharge between 1994 and 1998 accompanies an absence of data, not necessarily an absence of blooms). As described by the models, river discharge, through concentrated low-flow periods and 'load' events, may provide a eutrophic source of nitrate conducive to seasonal bloom formation, while allaying immediate bloom formation during periods of peak discharge.

Although not observed in the modeling dataset compiled here, one independent study recently reported 'a shift in toxin-producing species associated with an overall restructuring of the phytoplankton community' for Monterey Bay (Jester et al. 2009). Jester et al. (2009) used a similar dataset to ours (Monterey Bay, 2000 to 2006); the discrepancy in *Pseudo-nitzschia* abundance observations between the datasets may be due to differences in spatial coverage. The 'shift in toxin-producing species' was defined by a sharp decline in the incidence of toxigenic *Pseudo-nitzschia* in the summer of 2004, which persisted until the end of the study in 2006. This period was marked by anomalously low upwelling conditions and anomalously

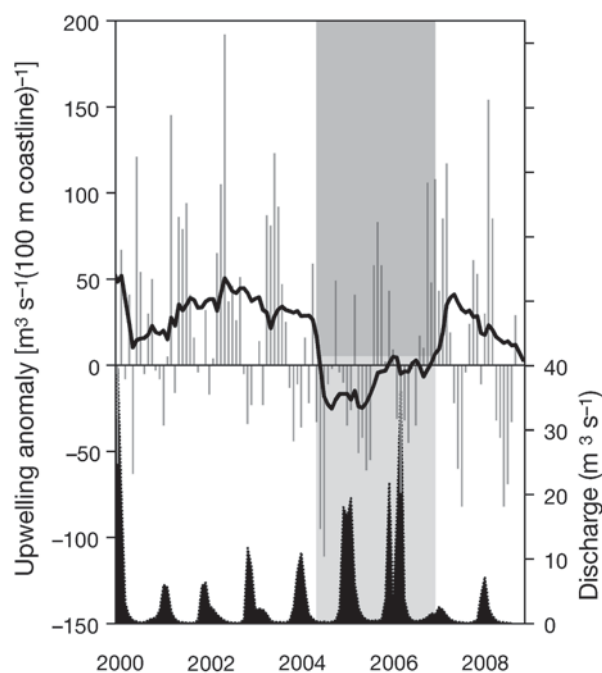


Fig. 6. Monthly upwelling anomaly for 36°N, 122°W (grey bars) with a 12 mo moving average trendline (black line), and the monthly mean discharge for the Pajaro River (area plot) are shown for the period from 2000 to 2008. An independent study addressing toxigenic *Pseudo-nitzschia* abundance in Monterey Bay for the period from 2000 to 2006 identified the summer 2004 period as a harmful species shifting-point (from *Pseudo-nitzschia* to *Alexandrium* and *Dinophysis*), and showed the summers from 2004 through 2006 to be periods of severely decreased *Pseudo-nitzschia* abundance (greyed area) (Jester et al. 2009). This 'shifted' period demonstrates relatively low upwelling and high river discharge activity

high periods of river discharge (Fig. 6), conditions which our models identify as non-conducive to toxigenic *Pseudo-nitzschia* blooms. Both of these conditions were alleviated in 2007, marked by a significant toxigenic *Pseudo-nitzschia* bloom event in Monterey Bay (Jester et al. 2009); a bloom of toxigenic *Pseudo-nitzschia* was also observed in 2008 (data not shown). According to the models and these observations, climatological conditions associated with low upwelling and high river discharge conditions may be conducive to suppressed toxigenic *Pseudo-nitzschia* bloom activity. The forecast of these conditions may now translate into the anticipation of large-scale shifts, such as the 'shift in toxin-producing species' described by Jester et al. (2009).

Bloom modeling versus toxin modeling

The monitoring of domoic acid for public health purposes is carried out continually by the CDPH and focuses, quite appropriately, on the protection of human health from domoic acid intoxication. This monitoring effort is more accurately described as the monitoring of domoic acid bioaccumulation in sentinel shellfish supplies (*Mytilus californianus*). Modeling or monitoring efforts that are focused on toxin load alone, while useful and appropriate for regulatory purposes, obviously do not allow for the estimation or monitoring of *Pseudo-nitzschia* blooms, which can be highly variable in their toxicity (Trainer et al. 2002, Marchetti et al. 2004, Anderson et al. 2006). This variability translates into a weak relationship between toxin bioaccumulation and toxigenic *Pseudo-nitzschia* abundance, evidenced here by CDPH/Cal-PreEMPT project data compiled from study sites in northern, central, and southern California over a 3 yr time period (Fig. 7). Note that there were cases where extreme bloom concentrations of toxigenic *Pseudo-nitzschia* were associated with sub-regulatory toxin levels ($<20 \mu\text{g g}^{-1}$), but no observed cases where sub-bloom concentrations of toxigenic *Pseudo-nitzschia* were associated with toxin levels approaching the regulatory limit in shellfish. Logistic regression models developed for toxigenic *Pseudo-nitzschia* blooms can therefore be used for detection of both acute and sub-acute toxic bloom events, while models developed for domoic acid alone will fail to address the injection of toxin into the system via sub-acute bloom events. This is a significant failure inherent to all toxin models, since chronic or early life stage exposure to sub-lethal levels of domoic acid are increasingly being recognized as an emerging threat to both human health and wildlife (Kreuder et al. 2005, Goldstein et al. 2008, Grattan et al. 2008, Ramsdell & Zabka 2008, M. Miller pers. comm.). By providing

estimations of all toxigenic *Pseudo-nitzschia* bloom events, whether low or high in toxicity, *Pseudo-nitzschia* bloom models have the unique ability to address this emerging threat. Ideally, future models should be developed for both cell abundance (present study) and for toxin production (Blum et al. 2006, Anderson et al. 2009). While the domoic acid data associated with the cases used herein were insufficient for inclusion of a toxin component, a 2-step model would maximize both regulatory monitoring and our understanding of the ecophysiological conditions associated with toxin production.

CONCLUSIONS

The models presented here demonstrate toxigenic *Pseudo-nitzschia* bloom classification rates of $\geq 75\%$. These predictive success rates are comparable to, or improved over, those reported for previous models of toxicity and generic *Pseudo-nitzschia* blooms. The assessment of our model alongside a chl *a* anomaly model, a useful tool designed for the detection of HABs more generally, demonstrates the capacity for improved predictive ability through more rigorous model development. Although we have reported the largest modeling dataset to date, consisting of 506 cases from 2002 to 2008, the removal of approximately 75% of the full dataset highlights the need for more consistent data collection. The parameters common to the 3

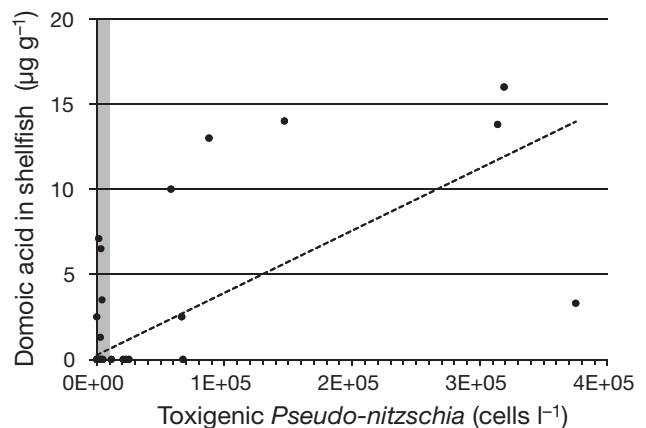


Fig. 7. Micrograms of domoic acid per gram shellfish plotted against counts of toxigenic *Pseudo-nitzschia*. Data are from study sites in northern, central, and southern California, monitored from 2005 to 2007 by Cal-PreEMPT in conjunction with the California Department of Public Health. Domoic acid never approached the regulatory limit of $20 \mu\text{g g}^{-1}$ shellfish when toxigenic *Pseudo-nitzschia* concentrations were at sub-bloom levels, i.e. $\leq 10,000 \text{ cells l}^{-1}$ (shaded area), and only began to approach the regulatory limit at cell concentrations that were significantly higher than $\leq 10,000 \text{ cells l}^{-1}$. The dashed line shows the results of a linear regression for domoic acid toxicity versus cell abundance

regional models developed for the west coast of the United States (Fig. 5A) provide direction for the design of ongoing and future HAB monitoring. We note that several parameters identified as being important (e.g. urea [Howard et al. 2007, Kudela et al. 2008a]; lithium [Subba Rao et al. 1998]; ammonium [Trainer et al. 2007]; iron and copper [Rue & Bruland 2001, Maldonado et al. 2002, Wells et al. 2005]) were not included for evaluation and could potentially further improve model prediction. All models included macronutrient variables as predictors of toxigenic *Pseudo-nitzschia* blooms, indicating the influence of upwelling and possibly cultural eutrophication on toxigenic *Pseudo-nitzschia* bloom proliferation. The seasonal significance of river discharge during periods associated with weak upwelling suggests that both natural (upwelling) and cultural (freshwater discharge) eutrophication processes, and the timing and seasonality of these processes, are significant factors influencing toxigenic *Pseudo-nitzschia* bloom dynamics. Although our models are specific to Monterey Bay, we have identified several factors common to all 3 modeling efforts for *Pseudo-nitzschia*. Given appropriate validation data, we suggest that some variant of this reduced subset of environmental variables could be applied to other regions, particularly similar coastal upwelling systems where *Pseudo-nitzschia* is prevalent (e.g. the west coast of the United States and Baja, Mexico, the southern Benguela, and the Iberian peninsula; Bates et al. 1998, Kudela et al. 2005, Fawcett et al. 2007).

Acknowledgements. We thank P. E. Miller (University of California, Santa Cruz, UCSC) and the Cal-PreEMPT team for contributing to the modeling dataset. We gratefully acknowledge the significant contribution made by M. W. Silver (UCSC) to the CIMT dataset and thank CIMT participants for their efforts. We thank G. J. Smith and K. Hayashi for their contribution of the MLML dataset. Special thanks to G. W. Langlois (CDPH) for providing shellfish toxicity data and an independent *Pseudo-nitzschia* monitoring dataset. Three anonymous reviewers contributed significantly to the improvement of this manuscript. Partial funding was provided by NOAA MERHAB Award NA04NOS4780239 (Cal-PreEMPT), NOAA Award NA160C2936 (CIMT), and, as a fellowship (J.Q.L.), from an anonymous donor through the Center for the Dynamics and Evolution of the Land-Sea Interface (CDELSI). This contribution is part of the Global Ecology and Oceanography of Harmful Algal Blooms (GEOHAB) Core Research Project on Harmful Algal Blooms in Upwelling Systems, and is MERHAB (Monitoring and Event Response for Harmful Algal Blooms) Publication Number 70.

LITERATURE CITED

- Addison RF, Stewart JE (1989) Domoic acid and the eastern Canadian molluscan shellfish industry. *Aquaculture* 77: 263–269
- Allen I, Smyth T, Siddorn J, Holt M (2008) How well can we forecast high biomass algal bloom events in a eutrophic coastal sea? *Harmful Algae* 8:70–76
- Anderson DM, Glibert PM, Burkholder JM (2002) Harmful algal blooms and eutrophication: nutrient sources, composition, and consequences. *Estuaries* 25:704–726
- Anderson CR, Brzezinski MA, Washburn L, Kudela R (2006) Circulation and environmental conditions during a toxigenic *Pseudo-nitzschia* bloom in The Santa Barbara Channel, California. *Mar Ecol Prog Ser* 327:119–133
- Anderson CR, Siegel DA, Kudela RM, Brzezinski MA (2009) Empirical models of toxigenic *Pseudo-nitzschia* blooms: potential use as a remote detection tool in the Santa Barbara Channel. *Harmful Algae* 8:478–492
- Bates SS, Trainer VL (2006) The ecology of harmful diatoms. In: Granéli E, Turner J (eds) *Ecology of harmful algae*, Vol 189. Springer-Verlag, Heidelberg, p 81–93
- Bates SS, Bird CJ, de Freitas ASW, Foxall RA and others (1989) Pennate diatom *Nitzschia pungens* as the primary source of domoic acid, a toxin in shellfish from eastern Prince Edward Island, Canada. *Can J Fish Aquat Sci* 46: 1203–1215
- Bates SS, Garrison DL, Horner RA (1998) Bloom dynamics and physiology of domoic-acid-producing *Pseudo-nitzschia* species. In: Anderson D, Cembella A, Hallegraeff G (eds) *Physiological ecology of harmful algal blooms*. Springer-Verlag, Heidelberg, p 267–292
- Beltrán AS, Palafox-Urbe M, Grajales-Montiel J, Cruz-Villacorta A, Ochoa JL (1997) Sea bird mortality at Cabo San Lucas, Mexico: evidence that toxic diatom blooms are spreading. *Toxicon* 35:447–453
- Bird CJ, Wright JLC (1989) The shellfish toxin domoic acid. *World Aquac* 20:40–41
- Blauw AN, Anderson P, Estrada M, Johansen M and others (2006) The use of fuzzy logic for data analysis and modeling of European harmful algal blooms: results of the HABES project. *Afr J Mar Sci* 28:365–369
- Blum I, Subba Rao DV, Pan Y, Swaminathan S, Adams NG (2006) Development of statistical models for prediction of the neurotoxin domoic acid levels in the pennate diatom *Pseudonitzschia pungens* f. *multiseries* utilizing data from cultures and natural blooms. In: Subba Rao DV (ed) *Algal cultures, analogues of blooms and applications*. Science Publishers, Enfield, NH, p 891–916
- Buck KR, Uttal-Cooke L, Pilskaln CH, Roelke DL and others (1992) Autecology of the diatom *Pseudonitzschia australis*, a domoic acid producer, from Monterey Bay, California. *Mar Ecol Prog Ser* 84:293–302
- CCLEAN (Central Coast Long-Term Environmental Assessment Network) (2006) 2004–2005 annual report. Available at: www.cclean.org/ftp/CCLEAN%20Final%2004-05.pdf
- CCLEAN (Central Coast Long-Term Environmental Assessment Network) (2007) Program overview 2001–2006. Available at: www.cclean.org/ftp/CCLEAN%2001-06%20Overview%20copy.pdf
- Fan X, Wang L (1998) Comparing linear discriminant function with logistic regression for the two-group classification problem. Annual Meeting of the American Educational Research Association, San Diego, CA (April 13–17, 1998)
- Fawcett A, Pitcher GC, Bernard S, Cembella A, Kudela RM (2007) Contrasting wind patterns and toxigenic phytoplankton in the southern Benguela upwelling system. *Mar Ecol Prog Ser* 348:19–31
- Fehling J, Davidson K, Bolch C, Tett P (2006) Seasonality of *Pseudo-nitzschia* spp. (Bacillariophyceae) in western Scottish waters. *Mar Ecol Prog Ser* 323:91–105
- Fisher WS, Malone TC, Giattina JD (2003) A pilot project to detect and forecast harmful algal blooms in the northern Gulf of Mexico. *Environ Monit Assess* 81:373–381

- Fritz L, Quilliam MA, Wright JLC, Beale AM, Work TM (1992) An outbreak of domoic acid poisoning attributed to the pennate diatom *Pseudonitzschia australis*. *J Phycol* 28: 439–442
- Glibert PM, Anderson DM, Gentien P, Granéli E, Sellner KG (2005) The global, complex phenomena of harmful algal blooms. *Oceanography (Wash DC)* 18:136–147
- Goldberg JD (2003) Domoic acid in the benthic food web of Monterey Bay, California. Master's thesis. California State University Monterey Bay, Monterey, CA
- Goldstein T, Mazet JAK, Zabka TS, Langlois G and others (2008) Novel symptomatology and changing epidemiology of domoic acid toxicosis in California sea lions (*Zalophus californianus*): an increasing risk to marine mammal health. *Proc R Soc Lond B Biol Sci* 275:267–276
- Grattan LM, Roberts S, Trainer V, Boushey C and others (2008) Domoic acid neurotoxicity in Native Americans in the Pacific Northwest: human health project methods and update. In: *Proc 4th Symp on harmful algae in the US*. US National Office for Harmful Algal Blooms
- Hallegraeff GM (1993) A review of harmful algal blooms and their apparent global increase. *Phycologia* 32:79–99
- Hensher DA, Johnson LW (1981) Applied discrete-choice modeling. Croom Helm, London
- Howard MA, Cochlan WP, Ladizinsky N, Kudela RM (2007) Nitrogenous preference of toxigenic *Pseudo-nitzschia australis* (Bacillariophyceae) from field and laboratory experiments. *Harmful Algae* 6:206–217
- Jester R, Lefebvre K, Langlois G, Vigilant V, Baugh K, Silver MW (2009) A shift in the dominant toxin-producing algal species in central California alters phycotoxins in food webs. *Harmful Algae* 8:291–298
- Johnsen G, Sakshuag E (2000) Monitoring of harmful algal blooms along the Norwegian coast using bio-optical methods. *S Afr J Mar Sci* 22:309–321
- Knepel K, Bogren K (2001) Determination of orthophosphorous by flow injection analysis in seawaters: QuickChem Method 31-113-01-1-H. In: *Saline methods of analysis*. Lachat Instruments, Milwaukee, WI
- Kreuder C, Miller MA, Lowenstine LJ, Conrad PA, Carpenter TE, Jessup DA, Mazet JA (2005) Evaluation of cardiac lesions and risk factors associated with myocarditis and dilated cardiomyopathy in southern sea otters (*Enhydra lutris nereis*). *Am J Vet Res* 66:289–299
- Kudela R, Pitcher G, Probyn T, Figueiras F, Moita T, Trainer V (2005) Harmful algal blooms in coastal upwelling systems. *Oceanography (Wash DC)* 18:184–197
- Kudela RM, Lane JQ, Cochlan WP (2008a) The potential role of anthropogenically derived nitrogen in the growth of harmful algae in California, USA. *Harmful Algae* 8: 103–110
- Kudela RM, Ryan JP, Blakely MD, Lane JQ, Peterson TD (2008b) Linking the physiology and ecology of *Cochlodinium* to better understand harmful algal bloom events: a comparative approach. *Harmful Algae* 7:278–292
- Lefebvre KA, Powell CL, Busman M, Doucette GJ and others (1999) Detection of domoic acid in northern anchovies and California sea lions associated with an unusual mortality event. *Nat Toxins* 7:85–92
- Lefebvre KA, Bargu S, Kieckhefer T, Silver MW (2002a) From sanddabs to blue whales: the pervasiveness of domoic acid. *Toxicon* 40:971–977
- Lefebvre KA, Silver MW, Coale SL, Tjeerdema RS (2002b) Domoic acid in planktivorous fish in relation to toxic *Pseudo-nitzschia* cell densities. *Mar Biol* 140:625–631
- Maldonado MT, Hughes MP, Rue EL, Wells ML (2002) The effect of Fe and Cu on growth and domoic acid production by *Pseudo-nitzschia multiseries* and *Pseudo-nitzschia australis*. *Limnol Oceanogr* 47:515–526
- Marchetti A, Trainer VL, Harrison PJ (2004) Environmental conditions and phytoplankton dynamics associated with *Pseudo-nitzschia* abundance and domoic acid in the Juan de Fuca eddy. *Mar Ecol Prog Ser* 281:1–12
- McGillicuddy Jr D, Anderson D, Lynch D, Townsend D (2005) Mechanisms regulating large-scale seasonal fluctuations in *Alexandrium fundyense* populations in the Gulf of Maine: results from a physical–biological model. *Deep Sea Res II* 52:2698–2714
- McPhee-Shaw EE, Siegel DA, Washburn L, Brzezinski MA, Jones JL, Leydecker A, Melack J (2007) Mechanisms for nutrient delivery to the inner shelf: observations from the Santa Barbara Channel. *Limnol Oceanogr* 52: 1748–1766
- Menard S (1995) Applied logistic regression analysis. Sage, Thousand Oaks, CA
- Miller PE, Scholin CA (1998) Identification and enumeration of cultured and wild *Pseudo-nitzschia* (Bacillariophyceae) using species-specific LSU rRNA-targeted fluorescent probes and filter-based whole cell hybridization. *J Phycol* 34:371–382
- Neter J, Wasserman W, Kutner M (1989) Applied linear regression models. Irwin, Homewood, IL
- Pennington JT, Chavez FP (2000) Seasonal fluctuations of temperature, salinity, nitrate, chlorophyll and primary production at station H3/M1 over 1989–1996 in Monterey Bay, California. *Deep Sea Res II* 47:947–973
- Ramsdell JS, Zabka TS (2008) *In utero* domoic acid toxicity: a fetal basis to adult disease in the California sea lion (*Zalophus californianus*). *Mar Drugs* 6:262–290
- Rue E, Bruland K (2001) Domoic acid binds iron and copper: a possible role for the toxin produced by the marine diatom *Pseudo-nitzschia*. *Mar Chem* 76:127–134
- Ryan JP, Gower JFR, King SA, Bissett WP and others (2008) A coastal ocean extreme bloom incubator. *Geophys Res Lett* 35, L12602, doi:10.1029/2008GL034081
- Schofield O, Grzymiski J, Bissett WP, Kirkpatrick GJ, Millie DF, Moline MA, Roesler CS (1999) Optical monitoring and forecasting systems for harmful algal blooms: possibility of pipe dream? *J Phycol* 35:1477–1496
- Scholin CA, Gulland F, Doucette GJ, Benson S and others (2000) Mortality of sea lions along the central California coast linked to a toxic diatom bloom. *Nature* 403:80–84
- Smith P, Bogren K (2001a) Determination of nitrate and/or nitrite in brackish or seawater by flow injection analysis colorimeter: QuickChem Method 31-107-04-1-E. In: *Saline methods of analysis*. Lachat Instruments, Milwaukee, WI
- Smith P, Bogren K (2001b) Determination of silicate in brackish or seawater by flow injection analysis colorimeter: QuickChem Method 31-114-27-1-C. In: *Saline methods of analysis*. Lachat Instruments, Milwaukee, WI
- Stumpf R, Tomlinson M, Calkins J, Kirkpatrick B and others (2009) Skill assessment for an operational algal bloom forecast system. *J Mar Syst* 76:151–161
- Subba Rao DVS, Pan Y, Mukhida K (1998) Production of domoic acid by *Pseudo-nitzschia multiseries* Hasle, affected by lithium. *PSZN I: Mar Ecol* 19:31–36
- Tomlinson MC, Stumpf RP, Ranisbrahmanakul V, Truby EW and others (2004) Evaluation of the use of SeaWiFS imagery for detecting *Karenia brevis* harmful algal blooms in the eastern Gulf of Mexico. *Remote Sens Environ* 91: 293–303
- Trainer VL, Adams NG, Bill BD, Anulacion BF, Wekell JC (1998) Concentration and dispersal of a *Pseudo-nitzschia*

- bloom in Penn Cove, Washington, USA. *Nat Toxins* 6: 113–126
- Trainer VL, Adams NG, Bill BD, Stehr CM and others (2000) Domoic acid production near California coastal upwelling zones, June 1998. *Limnol Oceanogr* 45:1818–1833
- Trainer VL, Adams NG, Wekell JC (2001) Domoic acid producing *Pseudo-nitzschia* species off the U.S. west coast associated with toxification events. In: Hallegraeff GM, Blackburn SI, Bolch CJ, Lewis RJ (eds) Harmful algal blooms 2000. Intergovernmental Oceanographic Commission of UNESCO, Paris, p 46–49
- Trainer VL, Hickey BM, Horner RA (2002) Biological and physical dynamics of domoic acid production off the Washington coast. *Limnol Oceanogr* 47:1438–1446
- Trainer VL, Cochlan WP, Erickson A, Bill BD, Cox FH, Borchert JA, Lefebvre KA (2007) Recent domoic acid closures of shellfish harvest areas in Washington State inland waterways. *Harmful Algae* 6:449–459
- Villac MC (1996) Synecology of the genus *Pseudo-nitzschia* H. Peragallo from Monterey Bay, California, USA. PhD dissertation, Texas A&M University, College Station
- Walz PM (1995) *Pseudo-nitzschia* species and domoic acid in Monterey Bay, CA. PhD dissertation, University of California, Santa Cruz
- Walz PM, Garrison DL, Graham WM, Cattet MA, Tjeerdema RS, Silver MW (1994) Domoic acid-producing diatom blooms in the Monterey Bay, California: 1991–1993. *Nat Toxins* 2:271–279
- Warrick JA, Mertes LAK, Siegel DA, Mackenzie C (2004) Estimating suspended sediment concentrations in turbid coastal waters of the Santa Barbara Channel with SeaWiFS. *Int J Remote Sens* 25:1995–2002
- Warrick JA, DiGiacomo PM, Weisberg SB, Nezlin NP and others (2007) River plume patterns and dynamics within the Southern California Bight. *Cont Shelf Res* 27:2427–2448
- Wells ML, Trick CG, Cochlan WP, Hughes MP, Trainer VL (2005) Domoic acid: the synergy of iron, copper, and the toxicity of diatoms. *Limnol Oceanogr* 50:1908–1917
- Welschmeyer NA (1994) Fluorometric analysis of chlorophyll *a* in the presence of chlorophyll *b* and pheopigments. *Limnol Oceanogr* 39:1985–1992
- Work TM, Beale AM, Fritz L, Quilliam MA, Silver M, Buck K, Wright JLC (1993) Domoic acid intoxication of brown pelicans and cormorants in Santa Cruz, California. In: Smayda TJ, Shimizu Y (eds) Toxic phytoplankton blooms in the sea. Elsevier Science Publication B.V., Amsterdam, p 643–650
- Wynne T, Stumpf R, Richardson A (2006) Discerning resuspended chlorophyll concentrations from ocean color satellite imagery. *Cont Shelf Res* 26:2583–2597

Editorial responsibility: Matthias Seaman, Oldendorf/Luhe, Germany

*Submitted: September 25, 2008; Accepted: March 5, 2009
Proofs received from author(s): April 21, 2009*