AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2014GL062937

Key Point:
- Global estimation of oceanic nitrate variability through SST, Chl, and MLD

Supporting Information:
- Readme
- Data Set S1
- Figure S1

Correspondence to:
L. Arteaga,
larteaga@geomar.de

Citation:
Arteaga, L., M. Pahlow, and A. Oschlies (2015), Global monthly sea surface nitrate fields estimated from remotely sensed sea surface temperature, chlorophyll, and modeled mixed layer depth, *Geophys. Res. Lett.*, *42*, 1130–1138, doi:10.1002/2014GL062937.

Received 19 DEC 2014
Accepted 12 JAN 2015
Accepted article online 16 JAN 2015
Published online 23 FEB 2015

Global monthly sea surface nitrate fields estimated from remotely sensed sea surface temperature, chlorophyll, and modeled mixed layer depth

Lionel Arteaga[1], Markus Pahlow[1], and Andreas Oschlies[1]

[1]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

Abstract Information about oceanic nitrate is crucial for making inferences about marine biological production and the efficiency of the biological carbon pump. While there are no optical properties that allow direct estimation of inorganic nitrogen, its correlation with other biogeochemical variables may permit its inference from satellite data. Here we report a new method for estimating monthly mean surface nitrate concentrations employing local multiple linear regressions on a global 1° by 1° resolution grid, using satellite-derived sea surface temperature, chlorophyll, and modeled mixed layer depth. Our method is able to reproduce the interannual variability of independent in situ nitrate observations at the Bermuda Atlantic Time Series, the Hawaii Ocean Time series, the California coast, and the southern New Zealand region. Our new method is shown to be more accurate than previous algorithms and thus can provide improved information on temporal and spatial nutrient variations beyond the climatological mean at regional and global scales.

1. Introduction

Monitoring and estimating primary production (PP) in marine environments rely heavily on satellite ocean-color observations due to their tremendously high spatial and temporal coverage not reached by any other current observing system. Many ocean color-based models that estimate PP depend on variables that can be retrieved from satellite observations using empirically derived algorithms (e.g., sea surface temperature (SST), photosynthetically active radiation, and chlorophyll (Chl)). PP is also estimated using ecosystem models, which often require the formulation of interactions among various agents, such as phytoplankton, zooplankton, and inorganic nutrients, representing the main ecosystem functions [*Pahlow et al.*, 2008]. Mechanistic models of phytoplankton growth offer the potential to provide further understanding of the physiological principles that regulate phytoplankton growth rates [e.g., *Geider et al.*, 1998; *Pahlow et al.*, 2013]. However, this kind of model commonly requires information on the availability of inorganic nutrients. Remote sensing tools offer synoptic information at the global level beyond the mean climatological state of different oceanic variables. Nevertheless, global estimations of sea surface nutrient concentrations are difficult to obtain from remote sensing, as there are no optical properties that allow their direct inference from satellite observations.

First attempts to indirectly assess nutrient availability in the surface ocean were based on nutrient-temperature-density relationships [*Kamykowski and Zentara*, 1986; *Garside and Garside*, 1995]. Most of the efforts have been directed at estimating dissolved inorganic nitrogen, which is regarded as the most immediate limiting nutrient in the ocean [*Falkowski*, 1997; *Tyrrell*, 1999; *Moore et al.*, 2013; *Arteaga et al.*, 2014]. More recently, nitrate concentrations have been diagnosed using satellite or in situ data of SST and mixed layer depth (MLD) [*Steinhoff et al.*, 2010] or SST and Chl [*Goes et al.*, 2000, 2004]. *Goes et al.* [2000, 2004] employed global relationships between surface nitrate concentration and SST and Chl, which are applicable to specific ocean basins. The methods developed by *Kamykowski et al.* [2002] and *Switzer et al.* [2003] used regional relationships based on 10° by 10° averages but provided only a relative assessment of nitrate availability. While *Steinhoff et al.* [2010] used a much finer (1° by 1°) grid for deriving their model, their algorithm is of regional nature and is based on a single multiple-regression function for the North Atlantic.

Despite these previous attempts, there is not yet an established product of nitrate concentrations in the surface layer of the ocean derived from remotely sensed data. Here we present a simple method to estimate monthly surface nitrate concentrations on a global scale by employing local multiple linear regressions using SST and Chl data from the Moderate Resolution Imaging Spectroradiometer (MODIS) and modeled

MLD. A major difference to previous methods, which are of an either regional or indirect nature, is that our approach is inherently local, i.e., it is applied on the same 1° by 1° resolution grid as the data on which it is based and thereby avoids the loss of information associated with averaging on a coarser grid.

2. Methods

In order to estimate surface nitrate concentration, we employ local multiple linear regressions using SST, MLD, and surface Chl as predictors of nitrate. Global surface Chl concentrations and SST are obtained from MODIS available from The Ocean Color site (http://oceancolor.gsfc.nasa.gov). To assess the general robustness of our method, we pragmatically employ different MLD models for the calibration and prediction periods. As the aim of our method is to provide monthly estimates of ocean surface nitrate concentrations in a semioperational mode, for the prediction period we choose to employ always the most recent—and presumably most accurate—available MLD model output. All MLD model outputs are obtained from the Ocean Productivity site of Oregon State University (http://science.oregonstate.edu/ocean.productivity/index.php): Monthly MLD outputs for 2003–2004 are from the SODAS model [*Clancy and Sadler*, 1992]. For the period January–June 2005, monthly MLD data were produced by an Isothermal Layer Depth (ILD) model of the Thermal Ocean Prediction System (TOPS), which is a model of the Fleet Numerical Meteorology and Oceanography Center (FNMOC), Monterey, California [*Clancy and Martin*, 1981; *Clancy and Pollak*, 1983; *Clancy and Sadler*, 1992] (ILD-TOPS). MLD for the period July 2005 to Sept 2008 is from a FNMOC high-resolution MLD criteria model. Finally, monthly MLD between October 2008 and December 2010 is from the Hybrid Coordinate Ocean Model [*Bleck*, 2002].

Since a sufficient seasonal coverage of remotely sensed Chl data could be obtained only for latitudes between 45°N and 45°S, we use only SST and MLD as predictors for higher latitudes. The coefficients of the linear regression are obtained for each 1° by 1° grid cell using the World Ocean Atlas 2009 (WOA09) monthly interpolated climatology of nitrate, and a training climatological data set of SST, Chl, and MLD. WOA09 is the only globally available training data set, and its temporal resolution is restricted to the monthly scale. Thus, it is currently not feasible to use a higher temporal resolution at the global scale. The training climatology is composed of average monthly SST and Chl MODIS data, and modeled MLD for January 2003 to December 2004:

$$\left(\beta_{SST}, \beta_{MLD}, \beta_{Chl}, C\right) = \text{regress}\left(\text{Nit}_{WOA09}, \text{SST}_{03-04}, \text{MLD}_{03-04}, \text{Chl}_{03-04}\right) \quad (1)$$

where β_{SST}, β_{MLD}, and β_{Chl} are the coefficients produced by the linear regression (regress) for monthly SST, MLD, and Chl respectively, and C is a local constant for each 1° by 1° grid cell (available as supporting information to this paper). Nit_{WOA09}, SST_{03-04}, MLD_{03-04}, and Chl_{03-04} are the monthly mean data sets employed to obtain the local regression coefficients.

Once the coefficients of the linear regression for each 1° by 1° grid cell are obtained, we force the multilinear model with an independent monthly data set of MODIS SST, Chl, and modeled MLD obtained for the period between January 2005 and December 2010:

$$\text{Nit}^{est} = \beta_{SST} \cdot \text{SST}_{05-10} + \beta_{MLD} \cdot \text{MLD}_{05-10} + \beta_{Chl} \cdot \text{Chl}_{05-10} + C \quad (2)$$

where Nit^{est} is the estimated nitrate. As a result, monthly global maps of estimated surface nitrate concentration are obtained on a 1° by 1° resolution grid for the period 2005–2010.

3. Results and Discussion

3.1. Global Patterns

Our predicted 1° by 1° temporally averaged surface nitrate concentration map for the period 2005–2010 is shown in Figure 1a. The spatial patterns of our modeled nitrate are consistent with the WOA09 nitrate climatology (Figure 1b). High nitrate concentrations are obtained at high latitudes, particularly the Southern Ocean, while tropical and subtropical areas show low concentrations, with the exception of the eastern Equatorial Pacific. The similarity of the spatial patterns of surface nitrate is encouraging given that the data set used to obtain the nitrate predictions (January 2005–December 2010) is independent from the data set used to derive the regression coefficients of the linear model (January 2003 to December 2004), and since we implicitly assume the general surface nitrate distribution in the global ocean for 2005–2010 to be similar to the climatology (WOA09).

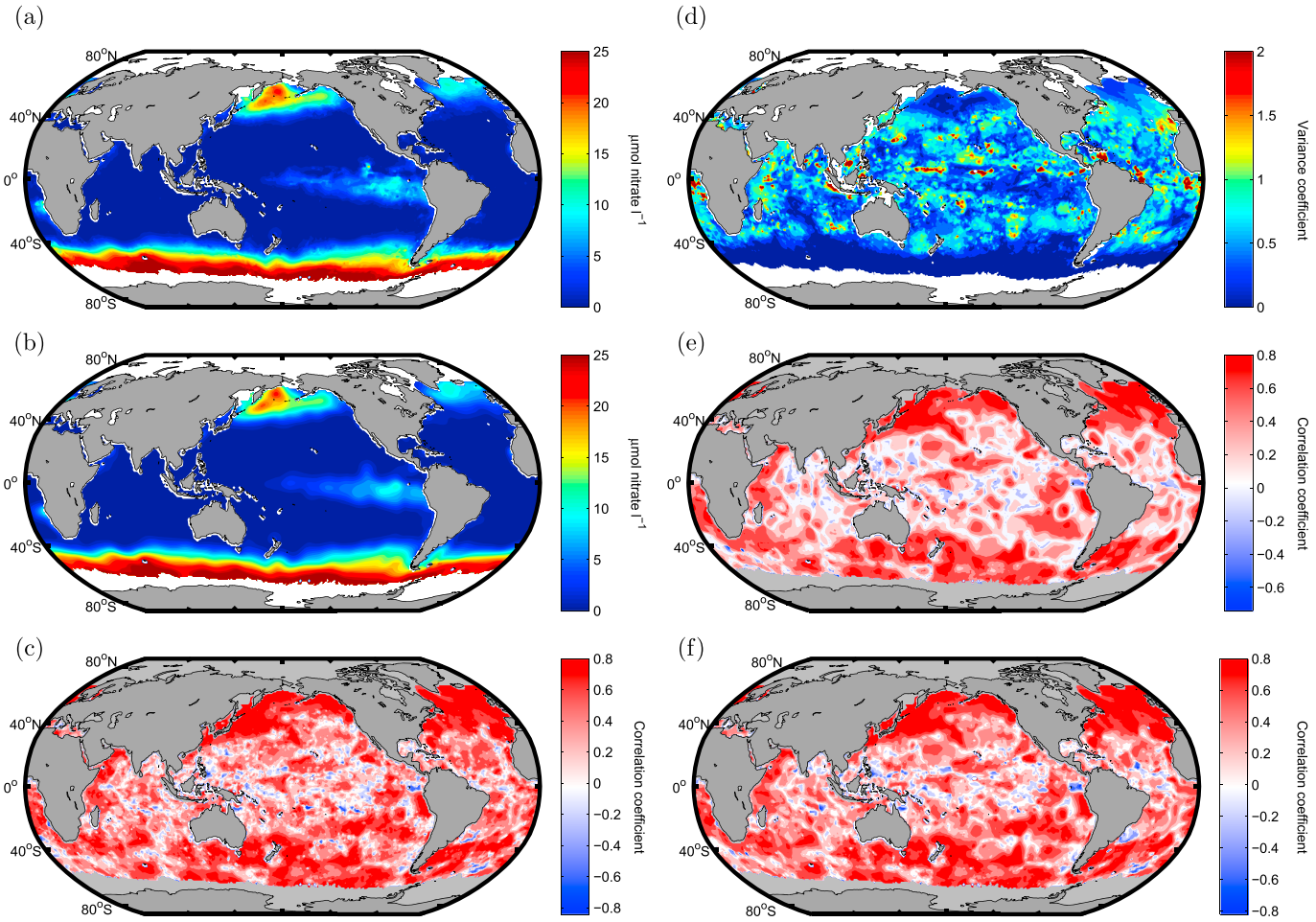

Figure 1. (a) Global surface nitrate concentration average for the period 2005–2010, estimated from SST, Chl, and MLD. (b) Global surface nitrate concentration climatology from World Ocean Atlas 2009 (WOA09). (c) Local correlation coefficient between our predicted nitrate concentrations and WOA09. (d) Variance coefficient of our 6 year predicted nitrate fields ($SD_{monthly05-10}/mean_{monthly05-10}$). (e) Local correlation coefficient between WOA09 and our predicted nitrate concentrations estimated from SST. (f) Local correlation coefficient between WOA09 and our predicted nitrate concentrations estimated from SST and MLD.

The local correlation coefficients (r) between our predicted monthly mean nitrate concentrations and WOA09 are highest at high latitudes, particularly in the northern North Atlantic and North Pacific, where r is close to 0.8 (Figure 1c). The correlation decreases toward the tropics, with very few scattered areas showing a negative correlation. The relative local monthly variance for our 6 years of predicted nitrate fields is calculated as the standard deviation of the whole monthly time series divided by its mean at each 1° by 1° grid point (Variance coefficient vc = $SD_{monthly05-10}/mean_{monthly05-10}$, Figure 1d). This relative temporal variance is highest in tropical regions (vc $\approx$ 2) and decreases at midlatitudes (vc $\approx$ 1), and toward the poles (vc $\approx$ 0).

The relation between SST, MLD, Chl, and nitrate is not the same at all points of the ocean. Seasonal variations in stratification, and the concomitant changes in SST and MLD, are much stronger at higher latitudes in comparison with middle and tropical latitudes. Thus, biotic factors can potentially have a stronger influence on nitrate availability in low latitudes. In order to analyse the effect of including or excluding our individual predictor variables, we also performed univariate (one factor) and bivariate (two factors) linear regressions to evaluate the predictive capacity of employing only SST and/or MLD. Interestingly, when compared against the climatology, nitrate estimated only from SST as a predictor shows a relatively high global pointwise correlation coefficient ($r = 0.46$) (Figure 1e) compared to the multiple linear regression shown above (including SST, MLD, and Chl, $r = 0.51$, Figure 1c). The bivariate regression using both SST and MLD as predictors results in a global average r of 0.51 (Figure 1f), which is the same as for the regression including also Chl. The lack of improvement of the global correlation when Chl is added as a factor could be due to the

small magnitude of the variations in monthly surface nitrate concentrations over tropical and subtropical regions relative to the average concentrations. Nevertheless, Chl appears to contribute substantially to the predictive power of our method with respect to interannual variations, e.g., revealed in time series observations, as we discuss below.

3.2. Interannual and Seasonal Variabilities

The fact that our results replicate the spatial and monthly patterns of the climatology quite well is an encouraging result. However, as our aim is to obtain interannually varying rather than climatological mean monthly nitrate estimations, we compare our predicted monthly surface nitrate concentrations for 2005–2010 with available independent nitrate data obtained for the top 100 m of the ocean from the Bermuda Atlantic Time-series Study (BATS) (http://bats.bios.edu), the Hawaii Ocean Time series (HOT) (http://hahana.soest.hawaii.edu), the California Cooperative Oceanic Fisheries Investigations program (CalCOFI) (http://www.calcofi.org), and the Munida Time Series in the South Pacific Ocean, maintained by the National Institute of Water and Atmospheric Research of New Zealand (http://cdiac.ornl.gov/oceans/Moorings/Munida.html), for the same period. Our predicted nitrate estimations are obtained for 1° by 1° grid boxes at 29°N, 60°W for BATS, 24°N, 158°W for HOT, 38°N, 123°W for CalCOFI, and 48°S, 173°E for Munida. Our 1° by 1° estimates agree favorably with the interannual variability of the four ocean time series (Figures 2a–2d). The seasonal variability is predicted to a somewhat lesser degree by the model. For HOT and BATS, estimated nitrate values are somewhat higher than the in situ data, while for CalCOFI the opposite occurs. This is partly due to differences in nitrate concentrations at the selected locations (HOT, BATS, and CalCOFI) between the climatological data (WOA09) used to obtain the linear regression coefficients and the in situ data (Figure 2). The correlation coefficient between model outputs and monthly averaged in situ data for 2005–2010 is relatively good for BATS ($r = 0.53$), CalCOFI ($r = 0.73$), and Munida ($r = 0.60$), but not very high for HOT ($r = 0.16$). The root-mean-square error (RMSE) is relatively high for BATS (0.24 μmol L^{-1}) and HOT (0.044 μmol L^{-1}), due to their inherently low average nitrate concentration. The RMSE for Munida is low (1.76 μmol L^{-1}) and similar to the RMSE obtained by *Sherlock et al.* [2007] with a basin-scale model for the same area. For CalCOFI the RMSE is also relatively low (2.50 μmol L^{-1}) and similar to that obtained by *Palacios et al.* [2013] for the northern region of the CalCOFI sampling program with a regional model.

The seasonality in the WOA09 climatology is markedly different from that in the in situ observations for BATS (Figure 2a). Nevertheless, the model output follows both the seasonal and interannual variability of the in situ data, despite the slight overestimation, which indicates a strong local coupling between SST, MLD, surface Chl, and nitrate, essentially allowing a reliable prediction of surface nitrate from satellite-derived observations of those variables. In particular, the extremely high nitrate concentration at BATS in the winter 2009–2010 is well captured by our regression approach (Figure 2a). This underlines the predictive capability of our method since these concentrations are well outside the range of nitrate concentrations in the training data set (Figure 2a).

The observed relationship between nitrate, MLD, and temperature is due to the vertical transport of nutrients and cold water caused by deepening of the surface mixed layer [*Garside and Garside*, 1995]. Nutrient utilization and depletion by phytoplankton is the main process responsible for the correlation between nitrate and surface Chl. The relative contribution of each of these processes on the control of surface nitrate variations could be considerably different in the distinct areas of the global ocean. A comparison of in situ nitrate observations with nitrate estimations using only MLD as a predictor shows that MLD by itself is a poor predictor of nitrate at BATS (Figure 3a, MLD). The same occurs when only SST is employed in the linear regression (Figure 3a, SST). However, using Chl as the only predictor (Figure 3a, Chl) yields the highest correlation with the observations ($r = 0.55$, Table 1). Furthermore, the correlation between observed and modeled nitrate at BATS is high and positive only when Chl is included as a variable in the regression analysis (Table 1). This points to the importance of biological controls of nitrate at BATS and probably oligotrophic regions in general. *Goes et al.* [2000] reached similar conclusions by including Chl in their model for the Pacific Ocean. While our model shows the highest r when Chl is used as a single factor, the high nitrate concentrations observed in the in situ data during 2009 are replicated only when SST and MLD are also included in the regression (Figure 2a). The importance of employing all three variables in the regression analysis is visually clear, however, this does not appear to be reflected in the RMSE, due to (slight) mismatches in the predicted and observed timing of the nitrate peak. When SSTs or Chls are used as the only predictors of nitrate at HOT, the correlation with the observations is negative (Figure 3b, SST and Chl). MLD as the only predictor results in a low positive correlation (Figure 3b, MLD). This suggests that nitrate

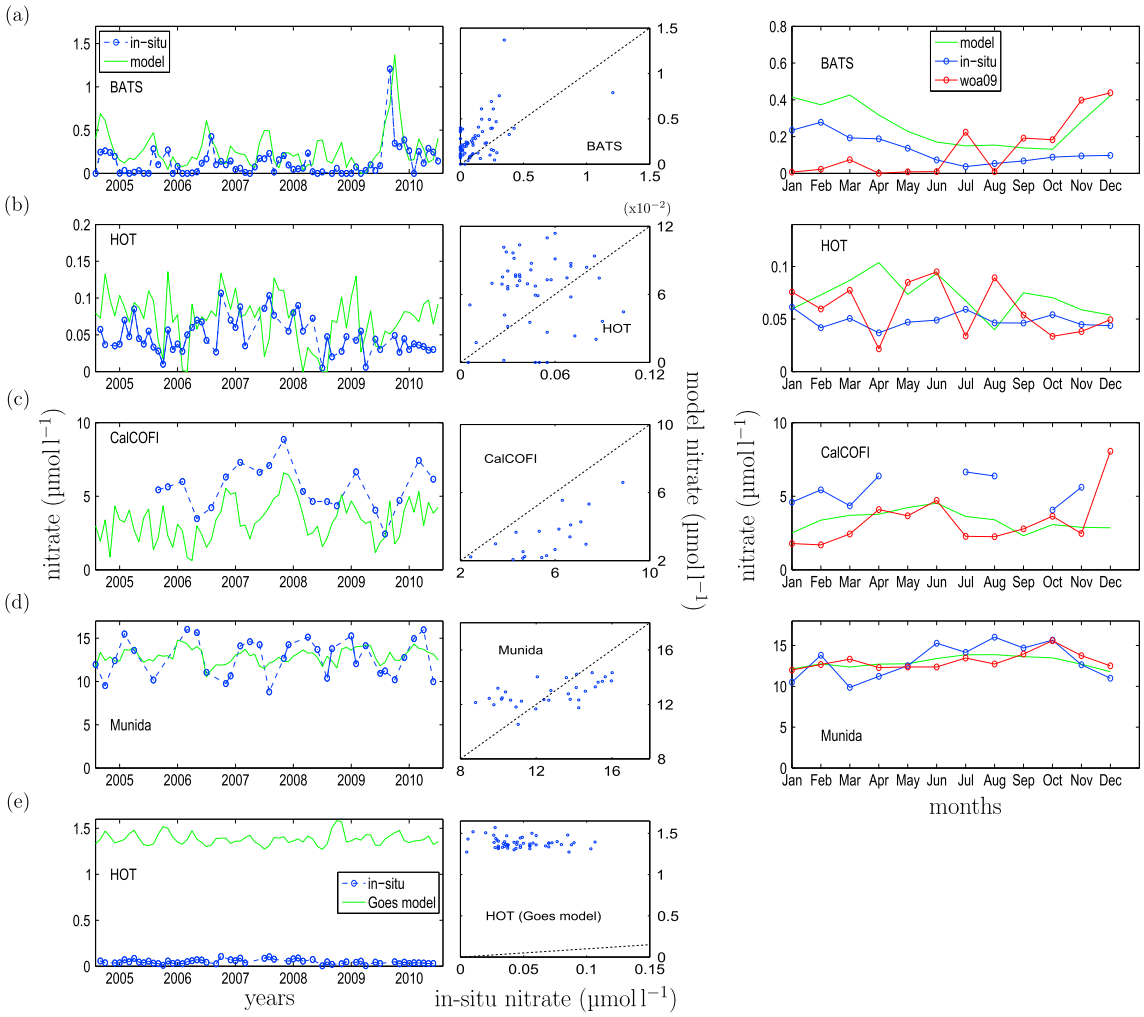

Figure 2. (left) Monthly predicted (green continuous line) and in situ measured (blue dashed line) surface nitrate concentrations for the period 2005–2010. (middle) Scatterplots for modeled and in situ nitrate concentrations and one-to-one (dashed) line. (right) Average monthly nitrate concentrations from January 2005 to December 2010. Model: green continuous line; In situ: blue dots connected with line; WOA09 climatology: red dots connected with line. Model predictions and observations are presented for (a) BATS, (b) HOT, (c) CalCOFI, and (d) Munida. (e) Nitrate prediction obtained for HOT with the model from *Goes et al.* [2000].

concentrations at HOT might also be strongly influenced by other physical or biogeochemical processes not captured by either SST, MLD, or Chl. While our approach is based on the idea that nitrate is either controlled by vertical mixing or phytoplankton consumption, other mechanisms, such as horizontal advection, have been suggested as important contributors of nutrient supply in the Pacific Ocean [*Dave and Lozier*, 2013].

SST and Chl are also poor single predictors of nitrate at CalCOFI. Similarly, as for HOT, MLD appears to be the best predictor in this region (Figure 3c). While combining MLD and SST results in the highest r for this region ($r = 0.77$, Table 1), the right magnitude of nitrate variations is only replicated when Chl is also included in the regression (Figure 2c). The obtained r for the regression with MLD, SST, and Chl ($r = 0.73$, Table 1) is only slightly lower than for only MLD and SST. *Palacios et al.* [2013] reached similar conclusions in their study, where the addition of physical and biological related variables such as temperature, salinity, and oxygen increased the explanatory power of their regional model for the CalCOFI area. Station Munida is located beyond 45°S; therefore, Chl cannot be included as part of the multiple regression. However, employing the available abiotic predictors SST and MLD individually results in high correlation coefficients (Table 1). For this station, using MLD or SST as the only predictor of nitrate results in a higher correlation coefficient than including both variables in the regression ($r = 0.75$ for MLD, $r = 0.66$ for SST, $r = 0.60$ for SST and MLD), which is likely due to a mismatch between WOA09 and in situ data at this site. As the difference between the

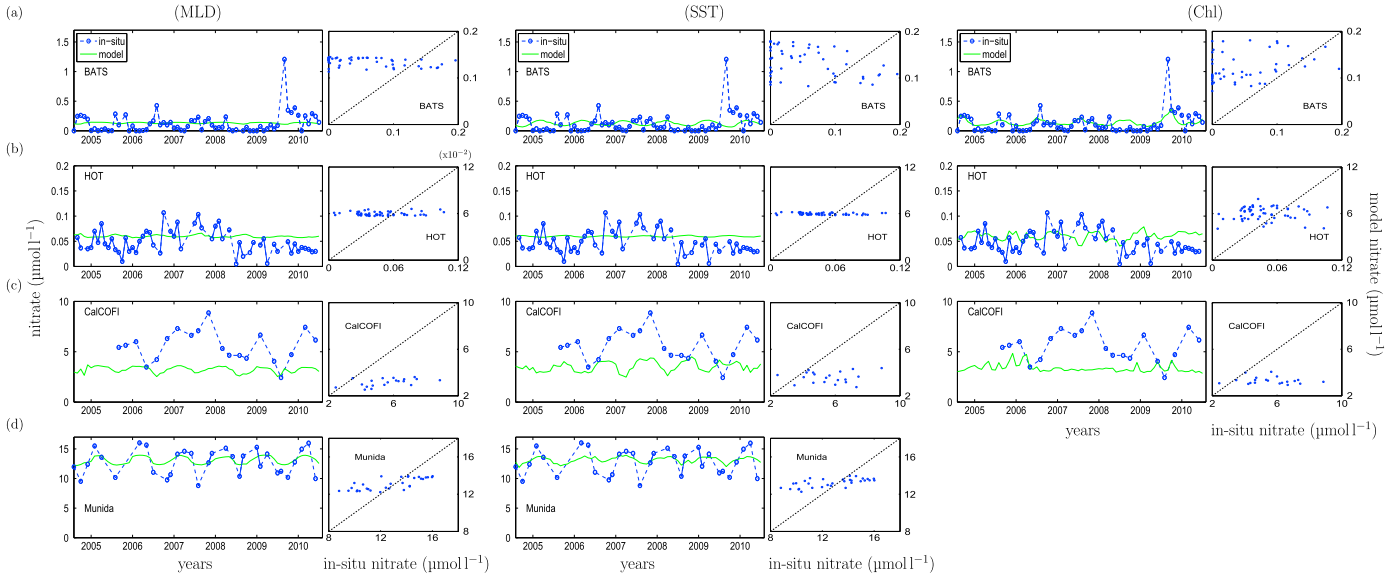

Figure 3. Monthly predicted (green continuous line) and in situ measured (blue dashed line) surface nitrate concentrations from January 2005 to December 2010 along with scatterplots for modeled and in situ nitrate concentrations and one-to-one (dashed) line, for (a) BATS, (b) HOT, (c) CalCOFI, and (d) Munida. Modeled nitrate was obtained with single linear regressions using only MLD (left), SST (middle), and Chl (right) as individual predictors.

correlations employing all available or individual variables is not very large, and it is impossible to discern beforehand which variable is the best predictor for nitrate in a particular point in the ocean, we thus stick to employing all three variables (SST, MLD, and Chl) in the regression analysis as long as they are available.

The use of a 1° by 1° grid avoids averaging and thereby minimizes information loss, which is a major distinction to most previous methods for estimating global nitrate distribution from satellite-derived data [e.g., *Goes et al.*, 2000]. As might be expected, our locally derived relations provide a much improved local predictive power compared to the basin-scale relationships. Figure 2e compares the model of *Goes et al.* [2000] for the nonequatorial Pacific using our forcing SST and Chl data set with observations at HOT. Although working well on the basin scale, the method of *Goes et al.* [2000] highly overestimates nitrate concentrations at HOT between 2005 and 2010, and the predicted nitrate is anticorrelated with the independent in situ observations (Figure 2e and Table 1).

3.3. Error Analysis

The predictive power of our method depends to a large extent on the accurate estimation of SST and Chl from remote sensing, and on the quality of the modeled MLD. Since we use linear regressions, systematic

Table 1. Correlation Coefficients (*r*) and Root-Mean-Square Errors (RMSE) for Nitrate Prediction Using Different Predictors (SST, MLD, and Chl) and In Situ Observations for 2005–2010 at HOT, BATS, CalCOFI, and Munida[a]

	r				RMSE			
Model variables	HOT	BATS	CalCOFI	Munida	HOT	BATS	CalCOFI	Munida
SST	−0.20	−0.43	0.07	0.66	0.025	0.19	2.50	1.94
MLD	0.10	−0.45	0.52	0.75	0.025	0.18	2.77	1.76
Chl	−0.002	0.55	-0.14	-	0.027	0.15	2.76	-
SST MLD	0.06	−0.36	0.77	0.60	0.025	0.18	2.50	1.76
SST Chl	−0.04	0.34	0.05	-	0.027	0.16	2.52	-
MLD Chl	0.08	0.52	0.27	-	0.037	0.22	2.81	-
SST MLD Chl	0.16	0.53	0.73	-	0.044	0.24	2.50	-
Goes model (HOT)	−0.16	-	-	-	1.34	-	-	-

[a]Goes model (HOT) refers to the model by *Goes et al.* [2000] evaluated against in situ data from HOT. Number of sample points for HOT = 55, BATS = 69, Munida = 31, and CalCOFI = 41.

errors have no effect. Thus, we restrict the error analysis to random errors (noise). To our knowledge, no method exists to separate systematic and random error components. Instead, we here employ the data used for validation to estimate the maximum random error which still allows reproducing the nitrate data in Figure 2. We apply various degrees of noise to the SST, MLD, and Chl data sets and then repeat the above training and validation procedure with the noisy data sets. The maximum noise level still reproducing the time series observations is then applied to the global ocean to obtain an estimate of the local uncertainty of the model prediction.

We introduce random noise to each of the forcing data sets (SST, MLD, and Chl) used in the linear regression, via,

$$\mathrm{ns} = \mathrm{data} + \sigma \cdot a \cdot x \tag{3}$$

where ns is the newly created "noisy signal," data is the original satellite-derived data (SST, Chl, or MLD for 2003–2004), σ is the temporal standard deviation of each grid point of the annual mean map of data, a is a coefficient determining the overall noise level, and x is a uniformly distributed random number between 0 and 1. We then compare the predictive power our method applied to the noisy data sets (i.e., $\mathrm{ns_{sst}}$, $\mathrm{ns_{chl}}$, $\mathrm{ns_{mld}}$), against that resulting from the original (unaltered) data sets.

We employ noise levels obtained with $a = 1$, 0.25, and 0.1. The main effect of the noise is a reduction in the predicted amplitude of the seasonal and interannual variations in nitrate concentrations. Only for the lowest noise level, i.e., $a = 0.1$, the resulting predicted nitrate is similar to the prediction without noise (Figure S1). Our method performs much worse with the noisy data sets with respect to the nitrate peak in early 2010, which is reproduced only with the lowest noise level (Figure S1b).

Our error analysis shows that the accuracy of our method can effectively be hampered by the presence of large random errors in the predictive data sets. The fact that the maximum random error compatible with the predictive power indicated in Figure 2 is obtained with $a = 0.1$ suggests a low random error in the satellite-based SST and Chl, and modeled MLD, which are used as predictors of nitrate in the linear regression. Keeping in mind that our artificial noise is added to the actual random error in the modeled and satellite-derived inputs, this result tells us that increasing the noise level (i.e., reducing the accuracy) by more than 10% over the current level would strongly interfere with our ability to obtain useful predictions of surface nitrate concentrations. While some uncertainty remains about the right magnitude of random errors in the forcing data sets, it appears unlikely that the actual noise level is much higher than our maximum estimate. For example, our relatively low upper limit of noise in the satellite-derived Chl compared to the previously estimated 35% [*Moore et al.*, 2009] suggests that either the local error in the satellite-derived and modeled inputs is lower than currently thought or that most of the error can be considered systematic at the local scale of our analysis.

To assess the global effect of our maximum error estimate, i.e., equivalent to adding 10 % of σ ($a = 0.1$) to any possibly existing noise, we compute the relative difference (reldiff) between the originally predicted nitrate and the nitrate obtained with a random error induced in the forcing satellite data sets with $a = 0.1$, on each 1° × 1° pixel (Figure S1d):

$$\mathrm{reldiff} = \frac{\mathrm{Nit}^{\mathrm{est}} - \mathrm{Nit}^{\mathrm{est}}_{a=0.1}}{0.5(\mathrm{Nit}^{\mathrm{est}} + \mathrm{Nit}^{\mathrm{est}}_{a=0.1})} \tag{4}$$

where $\mathrm{Nit}^{\mathrm{est}}$ is the originally predicted nitrate, $\mathrm{Nit}^{\mathrm{est}}_{a=0.1}$ is the nitrate predicted with the random error for $a = 0.1$ applied to the predictive data set. The errors reflect the combined effects of random errors in the three forcing data sets. The relative error is rather small and varies between ± 20%, with larger errors generally restricted to low and middle latitudes and very small errors (within ± 2%) found at high latitudes (Figure S1). Therefore, we conclude that the errors associated with the satellite-derived data sets do not appear to hamper our satellite-based method of predicting nitrate distributions for the world ocean.

4. Summary and Conclusions

We present a new method to estimate monthly surface nitrate distributions on a global 1° by 1° resolution grid. Our method is based on local multiple linear regressions applied at each 1° grid cell, employing SST, Chl, and MLD as predictors for nitrate. We evaluate the predictive power of our method against in situ data

from the oceanographic time series from stations BATS, HOT, CalCOFI, and Munida. Our modeled nitrate agrees well with the interannual variability of these stations, while the seasonal variability is reproduced to a lesser extent. Biological processes, represented by Chl, seem to have an important role controlling nitrate variations at BATS. While our analysis seems to indicate that the mechanisms controlling nitrate at HOT are not particularly well represented by either SST, MLD, or Chl, at Munida, nitrate can be well inferred through SST and MLD. At CalCOFI, MLD seems to be the main determinant of nitrate variations; however, SST and Chl are also required to reproduce the variability of the in situ observations. Despite a markedly different seasonality between the WOA09 nitrate and the in situ observations, our method is still able to follow the variability of the in situ data, indicating strong coupling between SST, MLD, surface Chl, and nitrate. Our error analysis of the predictive data set suggests that the method is robust as long as the errors in the forcing data sets do not exceed about 10% of the seasonal variance of the data. This provides some confidence in the use of satellite-derived SST and Chl and modeled MLD, to predict real surface nitrate concentrations in the global ocean.

Local and regional nitrate prediction models remain highly valuable to investigate controls of nitrate at the regional scale. Our nitrate prediction method allows to easily estimate nitrate variations at a global scale and offers the possibility to employ mechanistic models of marine biological production [e.g., *Pahlow et al.*, 2013] that require information on dissolved inorganic nitrogen availability. The sensitivity of phytoplankton growth rates to nutrient variations is an important constraint that needs to be taken into account in the monitoring and prediction of marine biological primary productivity under changing environmental conditions.

Acknowledgments

We wish to thank M. Behrenfeld and R. O'Malley for sharing their expertise on the mixed layer depth data obtained from the Ocean Productivity site of Oregon State University (http://www.science.oregonstate.edu/ocean.productivity/index.php). The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013, Space Theme, under grant agreement 282723 (OSS2015). This work is also a contribution of the Sonderforschungsbereich 754 "Biogeochemistry Interactions in the Tropical Ocean" (www.sfb754.de). All the details and web addresses needed to access the data used to produce our results are described in the section 2 of this paper, as well as in the supporting information.

The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

Arteaga, L., M. Pahlow, and A. Oschlies (2014), Global patterns of phytoplankton nutrient and light colimitation inferred from an optimality-based model, *Global Biogeochem. Cycles*, *28*(7), 648–661, doi:10.1002/2013GB004668.

Bleck, R. (2002), An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Model.*, *4*(1), 55–88, doi:10.1016/S1463-5003(01)00012-9.

Chen, D., L. M. Rothstein, and A. J. Busalacchi (1994), A hybrid vertical mixing scheme and its application to tropical ocean models, *J. Phys. Oceanogr.*, *24*(10), 2156–2179, doi:10.1175/1520-0485(1994)024<2156:AHVMSA>2.0.CO;2.

Clancy, R. M., and P. J. Martin (1981), Synoptic forecasting of the oceanic mixed layer using the navy's operational environmental data base: Present capabilities and future applications, *Bull. Am. Meteorol. Soc.*, *62*(6), 770, doi:10.1175/1520-0477(1981)062<0770:SFOTOM>2.0.CO;2.

Clancy, R. M., and K. D. Pollak (1983), A real-time synoptic ocean thermal analysis/forecast system, *Prog. Oceanogr.*, *12*(4), 383–424, doi:10.1016/0079-6611(83)90001-0.

Clancy, R. M., and W. D. Sadler (1992), The fleet numerical oceanography center suite of oceanographic models and products, *Weather Forecasting*, *7*(2), 307–327, doi:10.1175/1520-0434(1992)007<0307:TFNOCS>2.0.CO;2.

Dave, A. C., and M. S. Lozier (2013), Examining the global record of interannual variability in stratification and marine productivity in the low-latitude and mid-latitude ocean, *J. Geophy. Res. Oceans*, *118*, 3114–3127, doi:10.1002/jgrc.20224.

Falkowski, P. G. (1997), Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean, *Nature*, *387*, 272–275, doi:10.1038/246170a0.

Friedrichs, M. A., et al. (2009), Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean, *J. Mar. Syst.*, *76*, 113–133, doi:10.1016/j.jmarsys.2008.05.010.

Garside, C., and J. Garside (1995), Euphotic-zone nutrient algorithms for the NABE and EqPac study sites, *Deep Sea Res. Part II*, *42*(2–3), 335–347, doi:10.1016/0967-0645(95)00026-M.

Geider, R. J., H. L. MacIntyre, and T. Kana (1998), A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature, *Limnol. Oceanogr.*, *43*(4), 679–694, doi:10.4319/lo.1998.43.4.0679.

Goes, J. I., T. Saino, H. Oaku, J. Ishizaka, C. S. Wong, and Y. Nojiri (2000), Basin scale estimates of sea surface nitrate and new production from remotely sensed sea surface temperature and chlorophyll, *Geophys. Res. Lett.*, *27*(9), 1263–1266, doi:10.1029/1999GL002353.

Goes, J. I., H. D. R. Gomes, A. Limsakul, and T. Saino (2004), The influence of large-scale environmental changes on carbon export in the North Pacific Ocean using satellite and shipboard data, *Deep Sea Res., Part II*, *51*, 247–279, doi:10.1016/j.dsr2.2003.06.004.

Kamykowski, D., and S.-J. Zentara (1986), Predicting plant nutrient concentrations from temperature and sigma-t in the upper kilometer of the world ocean, *Deep Sea Res., Part A*, *33*(1), 89–105, doi:10.1016/0198-0149(86)90109-3.

Kamykowski, D., S.-J. Zentara, J. M. Morrison, and A. C. Switzer (2002), Dynamic global patterns of nitrate, phosphate, silicate, and iron availability and phytoplankton community composition from remote sensing data, *Global Biogeochem. Cycles*, *16*(4), 1077, doi:10.1029/2001GB001640.

Minnett, P. J., R. H. Evans, E. J. Kearns, and O. B. Brown (2002), Sea-surface temperature measured by the Moderate Resolution Imaging Spectroradiometer (MODIS), paper presented at IGARSS '02 Geoscience and Remote Sensing IEEE International Symposium, Anchorage, Alaska, 24–28 June 2002, vol. 2, pp. 1177–1179, doi:10.1109/IGARSS.2002.1025872.

Minnett, P. J., O. B. Brown, R. H. Evans, E. L. Key, E. J. Kearns, K. Kilpatrick, A. Kumar, K. A. Maillet, and G. Szczodrak (2004), Sea-surface temperature measurements from the Moderate-Resolution Imaging Spectroradiometer (MODIS) on Aqua and Terra, paper presented at IGARSS '04 Geoscience and Remote Sensing IEEE International Symposium, Anchorage, Alaska, 20–24 Sept. 2004, vol. 7, pp. 4576–4579, doi:10.1109/IGARSS.2004.1370173.

Moore, C. M., et al. (2013), Processes and patterns of oceanic nutrient limitation, *Nat. Geosci.*, *6*, 701–710, doi:10.1038/ngeo1765.

Moore, T. S., J. W. Campbell, and M. D. Dowell (2009), A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product, *Remote Sens. Environ.*, *113*, 2424–2430, doi:10.1016/j.rse.2009.07.016.

Palacios, D. M., E. L. Hazen, I. D. Schroeder, and S. J. Bograd (2013), Modeling the temperature-nitrate relationship in the coastal upwelling domain of the California current, *J. Geophys. Res. Oceans*, *118*, 3223–3239, doi:10.1002/jgrc.20216.

Pahlow, M., A. F. Vézina, B. Casault, H. Maass, L. Malloch, D. G. Wright, and Y. Lu (2008), Adaptive model of plankton dynamics for the North Atlantic, *Prog. Oceanogr.*, *76*(2), 151–191, doi:10.1016/j.pocean.2007.11.001.

Pahlow, M., H. Dietze, and A. Oschlies (2013), Optimality-based model of phytoplankton growth and diazotrophy, *Mar. Ecol. Prog. Ser.*, *489*, 1–16, doi:10.3354/meps10449.

Sherlock, V., S. Pickmere, K. Currie, M. Hadfield, S. Nodder, and P. W. Boyd (2007), Predictive accuracy of temperature-nitrate relationships for the oceanic mixed layer of the New Zealand region, *J. Geophys. Res.*, *112*, C06010, doi:10.1029/2006JC003562.

Steinhoff, T., T. Friedrich, S. E. Hartman, A. Oschlies, D. W. R. Wallace, and A. Körtzinger (2010), Estimating mixed layer nitrate in the North Atlantic Ocean, *Biogeosciences*, *7*(3), 795–807, doi:10.5194/bg-7-795-2010.

Switzer, A. C., D. Kamykowski, and S.-J. Zentara (2003), Mapping nitrate in the global ocean using remotely sensed sea surface temperature, *J. Geophys. Res.*, *108*(C8), 3280, doi:10.1029/2000JC000444.

Tyrrell, T. (1999), The relative influences of nitrogen and phosphorus on oceanic primary production, *Nature*, *400*(6744), 525–531, doi:10.1038/22941.