Global trends in marine nitrate N isotopes from observations and a neural network-based climatology

Patrick A. Rafter¹, Aaron Bagnell², Dario Marconi³, and Timothy DeVries²

¹: University of California, Irvine; ² University of California, Santa Barbara; ³: Princeton University

Abstract
Nitrate is a critical ingredient for life in the ocean because, as the most abundant form of fixed nitrogen in the ocean, it is an essential nutrient for primary production. The availability of marine nitrate is principally determined by biological processes, each having a distinct influence on the N isotopic composition of nitrate (nitrate \(\delta^{15}N\))—a property that informs much of our understanding of the marine N cycle as well as marine ecology, fisheries, and past ocean conditions. However, the sparse spatial distribution of nitrate \(\delta^{15}N\) observations makes it difficult to apply this useful property in global studies, or to facilitate robust model-data comparisons. Here, we use a compilation of published nitrate \(\delta^{15}N\) measurements \((n = 12277)\) and climatological maps of physical and biogeochemical tracers to create a surface-to-seafloor, 1° resolution map of nitrate \(\delta^{15}N\) using an Ensemble of Artificial Neural Networks (EANN). The strong correlation \((R = 0.87)\) and small mean difference \(<0.05\%)\) between EANN-estimated and observed nitrate \(\delta^{15}N\) indicates that the EANN provides a good estimate of climatological nitrate \(\delta^{15}N\) without a significant bias. The magnitude of observation-model residuals is consistent with the magnitude of seasonal-decadal changes in observed nitrate \(\delta^{15}N\) that are not captured by our climatological model. As such, these observation-constrained results provide a globally-resolved map of mean nitrate \(\delta^{15}N\) for observational and modeling studies of marine biogeochemistry, paleoceanography, and marine ecology.

1 Introduction
In contrast to other marine nutrients (e.g., phosphate and silicate), the inventory of nitrate \((NO_3^-)\) is mediated by biological processes, where the main source is \(N_2\) fixation by diazotrophic phytoplankton and the main sink is denitrification (via a microbial consortium in oxygen deficient waters and sediments) (Codispoti and Christensen, 1985). Biological processes also determine the distribution of marine nitrate throughout the water column, with phytoplankton assimilating nitrate / lowering nitrate concentrations in the surface ocean and the microbially-mediated degradation of organic matter in the subsurface. (The latter involving the multi-step process of ammonification (organic matter \(\rightarrow NH_4^+\)) and nitrification \((NH_4^+ \rightarrow NO_2^- \rightarrow NO_3^-)\).) By regulating the global inventory and distribution of marine nitrate, these N cycling processes control global net primary productivity, the transfer of nutrients to higher trophic levels such as fishes, and the strength of the ocean’s biological carbon pump (Dugdale and Goering, 1967).

Each of these biologically mediated N transformations affects the N isotopic composition of nitrate in unique ways (Fig.s 1A & 1B and see Section 2), adjusting the relative abundance of \(^{15}N\) and \(^{14}N\) in oceanic nitrate relative to the atmosphere. \(\delta^{15}N = \left(\frac{^{15}N}{^{14}N}_{sample} / \frac{^{15}N}{^{14}N}_{atmosphere}\right) - 1 \times 10^3\).
15N/14Nstandard) – 1), multiplied by 1000 to give units of per mil (%); see (Sigman and
Casciotti, 2001) for simplified equations from (Mariotti et al., 1981). Nitrate δ15N
measurements have become a powerful tool for understanding the ‘biogeochemical history’
of marine nitrate, which includes nitrate assimilation by phytoplankton (Miyake and Wada,
1967; Wada and Hattori, 1978), nitrogen fixation (Carpenter et al., 1997; Hoering & Ford,
1960), denitrification (Liu, 1979), and nitrification (Casciotti et al., 2013). For example, the
consumption of nitrate by denitrification (red line in Fig. 1A) has a larger impact on the
residual nitrate δ15N than does partial nitrate assimilation by phytoplankton (yellow line in
Fig. 1), and thus very high δ15N values serve as a fingerprint of denitrification. Nitrate δ15N
is also influenced by the addition of nitrate via remineralization of organic matter. The
exact influence of remineralization depends on the isotopic composition of the organic
matter, and could result in both higher or lower nitrate δ15N (Fig. 1A). Nitrate introduced
into the water column by the remineralization of organic matter formed by N2-fixing
phytoplankton has an isotopic composition close to that of air (0-1%), and serves to lower
the mean ocean δ15N (Fig. 1B). On the other hand, organic matter formed in regions where
the plankton use of most of the available nitrate can be isotopically heavy, and its
remineralization will increase the δ15N of ambient nitrate (Fig. 1B). The actual value of
organic matter δ15N is determined by: (1) the δ15N of nitrate delivered to the euphotic zone
(the subsurface source), which in turn is dependent on the degree of water-column
denitrification and (2) the degree of nitrate consumption at the ocean surface, with heavier
values associated with greater nitrate consumption (Fig. 1B). Accordingly, changes in
organic matter δ15N (and therefore sediment δ15N used for paleoceanographic work) can
reflect variability of the source nitrate δ15N and/or variability of the degree of nitrate
consumption (e.g., see (Rafter and Charles, 2012)).

Because of nitrate’s place at the base of the marine ecosystem, nitrate δ15N is also useful for
understanding the lifecycles of higher trophic level organisms such as fish (Graham et al.,
2007; Tawa et al., 2017) and fishery productivity (Finney et al., 2002, 2000). The δ15N of
whole sediment and microfossils provides insight by proxy of past ocean nitrate
transformations (Altabet and Francois, 1994; Galbraith et al., 2008; Kienast et al., 2008;
Rafter et al., 2012; Robinson et al., 2004; Sigman et al., 1999b)—work that places important
constraints on modern ocean N cycling (Altabet, 2007; Eugster et al., 2013; Ren et al.,
2017). With an understanding of the N transformations described above and their
influences on the N isotopic composition of nitrate, we can begin using nitrate δ15N
measurements to trace the integrated biogeochemical history of marine nitrate. However,
identifying basin- and global-scale trends in nitrate δ15N is challenged by the limited spatial
extent of nitrate δ15N observations (Fig. 2). Here, we compile a global database of nitrate
δ15N measurements (Fig. 2) and use an Ensemble Artificial Neural Network (EANN) to
produce a map of the global nitrate δ15N distribution at 1-degree spatial resolution. We find
that the mapped nitrate δ15N climatology matches the observations well and should be a
valuable tool for estimating mean conditions and for constraining predictive nitrate δ15N
models (Somes et al., 2010; Yang and Gruber, 2016). Below we briefly discuss how the
EANN was used to produce global maps of nitrate δ15N (Section 2), address the ability of
the EANN to match the measured δ15N (Section 3), and examine the EANN-mapped δ15N
climatology and global compilation of nitrate $\delta^{15}$N in the context of published work (Section 4).

2 Methods

2.1 Data Compilation

Nitrate $\delta^{15}$N observations (Fig. 2; references in Table 3) were compiled from studies dating from 1975 (Cline and Kaplan, 1975) to 2018 (Fripiat et al., 2018), including data from the GEOTRACES Intermediate Data Product (Schlitzer et al., 2018). Whenever possible, the data was acquired via the original author, but in other cases the data was estimated from the publication directly. All observations were treated equally, although the failure to remove nitrite when using the "denitrifier method" may bias the nitrate $\delta^{15}$N to low values (Rafter et al., 2013). These measurements have been identified as "nitrate+nitrite" in the dataset to acknowledge this potential biasing, which predominantly affects observations in the upper 100 m (Kemeny et al., 2016; Rafter et al., 2013).

2.2 Building the neural network model

We utilize an ensemble of artificial neural networks (EANNs) to interpolate our global ocean nitrate $\delta^{15}$N database (Fig. 2), producing complete 3D maps of the data. By utilizing an artificial neural network (ANN), a machine learning approach that effectively identifies nonlinear relationships between a target variable (the isotopic dataset) and a set of input features (other available ocean datasets), we can fill holes in our data sampling coverage of nitrate $\delta^{15}$N.

2.2.1 Binning target variables (Step 1)

We binned the nitrate $\delta^{15}$N observations (red symbols in Fig. 2) to the World Ocean Atlas 2009 (WOA09) grid with a 1-degree spatial resolution and 33 vertical depth layers (0-5500 m) (Garcia et al., 2010). When binning vertically, we use the midpoint between the depth values of one layer and the next as the partition between bins (e.g. the first depth layer has a value of 0 m, the second of 10 m, and the third of 20 m, so all nitrate isotopic data sampled between 0-5 m fall in the 0 m bin; between 5-15 m they fall in the 10 m bin, etc.). A point that lies right at the midpoint between depth intervals is binned to the shallower interval. If more than one raw data point falls in a grid cell we take the average of all those points as the value for that grid cell. Certain whole ship tracks of nitrate $\delta^{15}$N data were withheld from binning to be used as an independent validation set (see section 2.2.4).

2.2.2 Obtaining input features (Step 2)

Our input dataset contains a set of climatological values for physical and biogeochemical ocean parameters that form a non-linear relationship with the target data. These include objectively analyzed annual-mean fields for temperature, salinity, nitrate, oxygen, and phosphate taken from the WOA09 (https://www.nodc.noaa.gov/OC5/WOA09/woa09data.html) at 1-degree resolution. Additionally, daily chlorophyll data from Modis Aqua for the period Jan-1-2003 through Dec-31-2012 is binned to the WOA09 grid (as described in Step 1) to produce a climatological field of chlorophyll values, which we then log transform to reduce their dynamic range.
2.2.3 Training the ANN (Step 3)

The architecture of our ANN consists of a single hidden layer, containing 25 nodes, that connects the biological and physical input features (discussed in Step 2) to the target nitrate isotopic variable (as discussed in Step 1). The number of nodes in this hidden layer, as well as the number of input features, determines the number of adjustable weights (the free parameters) in the network. Because there is a danger of over-fitting the model, which occurs when the ANN is over-trained on a dataset so that it cannot generalize well when presented with new data, it is a good practice to have a large number of training data (≈7000 binned data points) relative to the number of weights (200 free parameters). A nonlinear activation function transforms the product of the weights and input features, creating the values assigned to nodes in the hidden layer. Our model utilizes the hyperbolic tangent as its activation function between the input and hidden layer as well as between the hidden and output layer due to its relative speed and general performance (Thimm and Fiesler, 1997). At the output layer, the network produces a prediction of the target nitrate isotopic data ($t_{\text{pred}}$), which it then compares to the actual values of that dataset ($t_{\text{data}}$). The ANN attempts to minimize the value of the cost function

$$\text{cost} = \frac{\sum_{i=1}^{n}(t_{\text{pred}}^{i} - t_{\text{data}}^{i})^2}{n}$$

by iteratively adjusting the weights using the Levenberg-Marquardt algorithm (Marquardt, 1963).

2.2.4 Validating the ANN (Step 4)

To ensure good generalization of the trained ANN to novel data, we randomly withhold 10% of the target isotopic data ($t_{\text{data}}$) to be used as an internal validation set for each network. This is data that the network never sees, meaning it does not factor into the cost function, so it works as a test of the ANN’s ability to generalize. This internal validation set acts as a gate-keeper to prevent poor models from being accepted into the ensemble of trained networks. Our pass criterion is an $R^2$ value greater than 0.9 between the ANN’s predicted value and the actual values of the validation set. A second, independent or ‘external’ validation set (blue symbols in Fig. 2), composed of complete ship transects from the high and low latitude ocean were omitted from binning in Step 1 and used to establish the performance of the entire ensemble. This independent validation set is never used in the process of developing our ensemble of ANNs.

2.2.5 Forming the Ensemble (Step 5)

The ensemble is formed by repeating Steps 3 to 4 until we obtain 25 trained networks for the nitrate $\delta^{15}$N dataset. Using an EANN instead of any single network provides several advantages. For example, the random initialization of the weight values in each network as well as differences in the training and internal validation sets used across members make it possible for many different networks to achieve similar performance on their respective validation set while generalizing areas with no data coverage differently. By performing this type of data subsampling and taking an ensemble average, similar to bootstrap
aggregating (Breiman, 1996) this approach improves the robustness of the generalization
in areas without data coverage, as demonstrated by the improved performance of the
ensemble versus any single member on the independent validation set. The range of values
given by the ensemble also provides a measure of the uncertainty for our estimations of
\( \delta^{15}N \).

### 3 Results

#### 3.1 Global nitrate \( \delta^{15}N \) observations

The global compilation of nitrate \( \delta^{15}N \) includes 1180 stations from all major ocean basins
and some minor seas (Fig. 2) giving a total of 12277 nitrate \( \delta^{15}N \) measurements. Within
this dataset, 1197 nitrate \( \delta^{15}N \) measurements were withheld from the EANN and used to
validate the EANN results to ensure good extrapolation (the ‘external’ validation dataset;
blue symbols in Fig. 2, see Section 2). With observations from the surface to as deep as
6002 m (Rafter et al., 2012), we find that nitrate \( \delta^{15}N \) ranges from \( \approx 1\% \) in the North
Atlantic (e.g., Marconi et al., 2015) to 68.7\% in the Eastern Tropical South Pacific
(Bourbonnais et al., 2015). Nitrate \( \delta^{15}N \) of \( \approx 1\% \) was also irregularly observed in the
shallow North and South Pacific (Liu et al., 1996; Yoshikawa et al., 2015). These latter
observations were included in the training dataset, although we should note that the
measurements using the ‘Devarda’s Alloy’ method (Liu et al., 1996) is thought to be biased
low (Altabet and Francois, 2001). Similarly, the inclusion of nitrate for ‘denitrifier method’
nitrate \( \delta^{15}N \) can bias the measurement to lower values (Kemeny et al., 2016; Rafter et al.,
2013).

#### 3.2 Marine nitrate \( \delta^{15}N \) observations-model comparison

The observed and EANN-predicted nitrate \( \delta^{15}N \) measurements are distributed around a 1:1
line in Fig. 3A (all data), with considerably less scatter for the deeper values (data >1000 m;
Fig. 3B). The correlation coefficient of determination for the observations versus the model
nitrate \( \delta^{15}N \) gives an \( R^2=0.75 \) for the observations used to train the EANN and an \( R^2 \) of 0.78
for the validation dataset. We can also examine the performance of the EANN with the
nitrate \( \delta^{15}N \) “residual” or the difference between observed and modeled \( \delta^{15}N \), which
indicates a mean residual or ‘mean bias’ value of -0.03\% for the entire dataset and
+0.18\% for the validation dataset.

Examining the observation-EANN residuals via the Root Mean Square Error (RMSE), we
find an RMSE of 1.94\% for the data used to train the EANN and an RMSE of 1.26\% for the
external validation dataset. There is a clear relationship between RMSE and depth, with a
significantly higher RMSE for the upper 500 m (Figs. 3C and 3D). Comparing these residual
values with dissolved oxygen concentrations (color in Fig. 3C), we find that >2\% RMSE for
the surface is associated with high oxygen while >2.7\% RMSE at \( \approx 250 \) m is associated with
the lowest oxygen. Furthermore, the RMSE of the observation-EANN residuals differs
between the datasets used to train the model (solid red line in Fig. 3D) and validate the
model (dashed line in Fig. 3D).

The RMSE patterns in Figs. 3C and 3D are to be expected given the natural variability in
nitrate \( \delta^{15}N \) driven by assimilation in the upper ocean and denitrification in the shallow
sub-surface—variability which is not captured by the climatological EANN. Rafter and Sigman (2016), presented a 5-year time-series of nitrate $\delta^{15}$N from the eastern equatorial Pacific, which showed that variability of nitrate assimilation produces seasonal-to-interannual deviations of $\delta^{15}$N of $\pm 2.5\%$, which is similar to the magnitude of the RMSE in the surface ocean ($2.2\%$). Although there are no nitrate $\delta^{15}$N time-series measurements from the subsurface Oxygen Deficient Zone (ODZ) waters where denitrification occurs, nitrate $\delta^{15}$N in ODZs presumably have similar seasonal-to-interannual (or longer timescale) variability due to changes in the rate and extent of water column denitrification (Deutsch et al., 2011; Yang et al., 2017). For example, a larger degree of nitrate undergoing water column denitrification would explain the extreme $\delta^{15}$N values at the bottom right of Fig. 3A—observations that all come from the ODZ waters of the Eastern Tropical South Pacific (Bourbonnais et al., 2015; Cacchione et al., 2013; Rafter et al., 2012; Ryabenko et al., 2012). Some of these very high nitrate $\delta^{15}$N values are associated with nitrate concentrations $<1$ $\mu$mol kg$^{-1}$ (Bourbonnais et al., 2015), values much lower than within our climatology for the subsurface Eastern Tropical South Pacific. These values thus represent episodic denitrification events that the EANN will not be able to capture because it is trained on climatological data. In the deep ocean where temporal variability is negligible, the observation-EANN residuals of $0.2\%$ are the same magnitude as the $\delta^{15}$N analytical errors, further emphasizing the ability of the model to match climatological average conditions.

### 4 Discussion

The EANN’s skillful estimate of mean marine nitrate $\delta^{15}$N can be useful for all areas of research using this widely used geochemical measurement. The zonal average view of EANN nitrate $\delta^{15}$N for each major ocean basin (Fig. 4) includes statistics comparing the observations versus EANN results above and below 1000 m. These region-specific statistics indicated a weaker correlation between EANN and observed nitrate $\delta^{15}$N in the deep Atlantic and Southern Ocean, despite low RMSE and negligible mean bias. This weak correlation likely derives from the limited range of deep nitrate $\delta^{15}$N variability ($\pm 0.1\%$) in these basins (see Fig. 5D).

The nitrate $\delta^{15}$N sections in Fig. 4 also show elevated values for the low latitude, upper mesopelagic Pacific (Fig. 4A) and Indian Oceans (Fig. 4D) where water column denitrification raises the residual nitrate $\delta^{15}$N (Fig. 1A). Viewing this elevated nitrate $\delta^{15}$N at the 250 m depth horizon (Fig. 5) better reveals the spatial heterogeneity of the observations and EANN results. (It is because of this intra-basin heterogeneity, and the fact that many observations are biased towards the areas of denitrification, that we did not plot the observed nitrate $\delta^{15}$N within the zonally-averaged Fig. 4 views.) The EANN error for the Fig. 5 depth intervals (Figs. 5E-SH) is the standard deviation of the 25 ensemble members of the EANN and shows a decrease in ensemble variability with depth—a trend that is consistent with the overall decrease in observed nitrate $\delta^{15}$N variability with depth (Figs. 4 & 5).

Below we inspect the observed and EANN-predicted nitrate $\delta^{15}$N and discuss the consistency of these results with our understanding of published work. This analysis begins with the spatial distribution of nitrate delivered to the upper ocean. We then discuss the
impacts of upper ocean nitrate assimilation on organic matter $\delta^{15}N$ and consider the influence of organic matter remineralization on sub-surface nitrate.

### 4.1 Subsurface and surface nitrate $\delta^{15}N$

The nitrate $\delta^{15}N$ distribution at 250 m depth (Fig. 5B) offers a view of nitrate at a depth that is deeper than source waters in many ocean regions (e.g., 100 to 150 m in the equatorial Pacific (Rafter and Sigman, 2016)), but is negligibly influenced by nitrate assimilation, and therefore provides a qualitative view of spatial trends in nitrate delivered to the surface ocean. Nitrate $\delta^{15}N$ at this depth is highest in the North and South Eastern Tropical Pacific and Arabian Seas (Fig. 5B), due to the influence of water column denitrification in the ODZs in these regions (Altabet et al., 2012; Bourbonnais et al., 2015; Ryabenko et al., 2012), which preferentially uses the light isotope and leaves the residual nitrate enriched in $^{15}N$.

Lowest $\delta^{15}N$ values of sub-surface nitrate are found in the Southern Ocean and in the North Atlantic. The North Atlantic subtropical gyre in particular has the lowest $\delta^{15}N$ values in any basin (Fig. 5B; also see (Fawcett et al., 2011; Knapp et al., 2005, 2008)), which can be attributed to the remineralization of low-$\delta^{15}N$ organic matter originating from N$_2$-fixation, which has a $\delta^{15}N$ between 0 and -1‰ (similar to atmospheric N$_2$; see Fig. 1B (Carpenter et al., 1997; Hoering & Ford, 1960)). Prior work argues that this nitrate $\delta^{15}N$ lowering requires the bulk of Atlantic N$_2$-fixation ($\approx$90%) to occur in the tropics (Marconi et al., 2017) followed by the advection of remineralized nitrate to the North Atlantic. This contrasts with numerical models arguing for high N$_2$-fixation rates in the North Atlantic (Ko et al., 2018). Similar local minima of sub-surface $\delta^{15}N$ appear in all the sub-tropical gyres (Fig. 5B), which is consistent with observations (Casciotti et al., 2008; Yoshikawa et al., 2015) and presumably indicates the importance of N$_2$-fixation in these regions (Ko et al., 2018) and others. The N$_2$-fixation $\delta^{15}N$ signal in the Pacific Ocean is counteracted by the influence of water-column denitrification in that basin, which imparts a high $\delta^{15}N$ signal, but a local minimum in $\delta^{15}N$ can still be seen in the Pacific subtropical gyres (Fig. 4A).

Nitrate assimilation by phytoplankton in the upper ocean is influenced by both the subsurface source nitrate $\delta^{15}N$ and the degree of nitrate assimilation (Miyake and Wada, 1967; Wada and Hattori, 1978) (Fig. 1B). This gives the expectation that average nitrate $\delta^{15}N$ values for the upper 50 m (Fig. 5A) will be consistently higher than those at 250 m (Fig. 5B). However, the highest values in the upper 50 m are not found above the ODZ regions, but are on the edges of high nitrate concentration upwelling zones in the Southern Ocean, equatorial Pacific, and subarctic gyres (contours in Fig. 2). Circulation in these ‘edge’ regions allows for nitrate to be advected along the surface, lengthening its time in the surface ocean and allowing more utilization to elevate the residual nitrate $\delta^{15}N$ pool. In other words, the degree of nitrate utilization appears to play a more important role in determining surface nitrate $\delta^{15}N$ than the initial value. (This is not the case for the organic matter $\delta^{15}N$ produced from this nitrate, which will be discussed more below and in Fig. 1B.)
Despite our expectation of higher nitrate $\delta^{15}N$ in the upper 50 m versus 250 m (Figs. 5A vs. 5B), we identify two types of regions where this difference is negative (Fig. 6): above ODZ waters and in subtropical gyres. The explanation for the negative values above the ODZ regions is that the nitrate $\delta^{15}N$ at 250 m must be much higher than the nitrate $\delta^{15}N$ upwelled to the surface. This is consistent with elevated ODZ nitrate $\delta^{15}N$ having an indirect path to waters outside of ODZ regions (Peters et al., 2017; Rafter et al., 2013). The subtropical gyres also have modeled nitrate $\delta^{15}N$ in the upper 50 m that is less than 250 m, but this finding is difficult to test with observations because of low nitrate concentrations. That said, the model predicts a lowering of the nitrate $\delta^{15}N$ in the upper ocean relative to the 250 m depths, which is consistent with N$_2$-fixation in these regions.

Our discussion above highlights the difficulty of distinguishing between the competing influences of the subsurface source nitrate $\delta^{15}N$ and the degree of nitrate utilization on residual nitrate $\delta^{15}N$. Clearly a static depth does not reflect the subsurface source of nitrate delivered to the surface and a more robust method for estimating this subsurface source needs to be developed. However, some generalizations can be made regarding the organic matter $\delta^{15}N$ produced in these regions and its potential influence (via remineralization) on subsurface nitrate throughout the water column via the export and remineralization of organic matter (Sigman et al., 2009). For example, a local minimum in $\delta^{15}N$ is visible at 250 m depth in the Eastern Equatorial Pacific (Fig. 5B; also discussed in several studies (Rafter et al., 2012; Rafter and Sigman, 2016)) is caused by the remineralization of organic matter with a low $\delta^{15}N$ due to partial nitrate consumption at the surface. Below we discuss these and other influences on intermediate-depth nitrate $\delta^{15}N$.

4.2 Intermediate-depth nitrate $\delta^{15}N$ variability

Waters at “intermediate” depths (here shown as the 750 m surface in Fig. 5C) are important because they are part of a large-scale circulation that initially upwells in the Southern Ocean and ultimately resupplies nutrients to the low latitude thermocline (Palter et al., 2010; Sarmiento et al., 2004; Toggweiler et al., 1991; Toggweiler and Carson, 1995). Within the context of this overturning, the nitrate upwelling in the Southern Ocean is initially $\approx5\%$ (Figs. 4C & 5C) and the $\delta^{15}N$ is elevated $\approx2\%$ by partial nitrate assimilation in surface waters as they are advected equatorward (see Figs. 5A and 6). Deep winter-time mixing in the Subantarctic Pacific converts these surface waters into mode and intermediate waters (Herraiz-Borreguero and Rintoul, 2011), introducing nitrate with a “pre-formed” $\delta^{15}N$ of $\approx6\%$ into the intermediate-depth South Pacific and South Atlantic (Rafter et al., 2012, 2013; Tuerena et al., 2015) at depths between $\approx600$-$1200$ m. The penetration of this pre-formed signal (nitrate $\geq 6\%$) into the interior can be clearly seen in the Atlantic Ocean between $\approx40^\circ$S to $20^\circ$N (Fig. 4B).

The same signal is carried with Southern Ocean mode and intermediate waters into the Pacific basin (Rafter et al., 2013), although it is difficult to distinguish in the model results against the higher background $\delta^{15}N$ in this basins (Figs. 4A, 4D, 5C). The same process presumably introduces elevated nitrate $\delta^{15}N$ to the Indian Ocean, which has similar values at this depth. Nitrate $\delta^{15}N$ increases from the Southern Ocean toward the equator in the Pacific and Indian Oceans, but not in the Atlantic (Fig. 5C). Organic matter has a lower $\delta^{15}N$
in the Atlantic than in the Pacific and Indian because of a lack of water-column

denitrification supplying high-\(\delta^{15}\)N water to the surface, and because of the high rates of

\(\text{N}_2\)-fixation which supply isotopically light N to organic matter (Marconi et al., 2017; Tuerena et al., 2015). This contrast in intermediate-depth nitrate \(\delta^{15}\)N can be traced to the

lower \(\delta^{15}\)N of organic matter remineralized in this region—an explanation that is also

consistent with enhanced \(\text{N}_2\) fixation in the tropical Atlantic (Marconi et al., 2017). The

increase in intermediate-depth nitrate \(\delta^{15}\)N from the Subantarctic to the tropical Pacific

appears from the remineralization of organic matter with a \(\delta^{15}\)N elevated by high

source nitrate \(\delta^{15}\)N (near the ODZ) or extreme elevation of residual nitrate \(\delta^{15}\)N (advected

along the surface away from the equator; see high surface nitrate \(\delta^{15}\)N in Fig. 5A). Previous

work suggests that direct mixing with denitrified waters represents only a small fraction of

the change from the pre-formed high latitude value (≈6‰) to tropical nitrate \(\delta^{15}\)N of \(\approx7\)‰

(Peters et al., 2017; Rafter et al., 2013; Rafter and Charles, 2012). Equivalent processes

must drive the \(\delta^{15}\)N in the intermediate-depth Indian Ocean, which is similarly elevated in

the EANN, although direct observations are needed in order to confirm how well the EANN

extrapolates in this region.

4.4 Deep-sea nitrate \(\delta^{15}\)N trends

Our discussion above suggests that the basin-scale balance of \(\text{N}_2\)-fixation and water-column

denitrification is a major contributor to inter-basin nitrate \(\delta^{15}\)N gradients in the upper

ocean, lowering values in the Atlantic Oceans compared to the Pacific and Indian Oceans.

Averaging EANN nitrate \(\delta^{15}\)N from the surface to 5500 m for each ocean basin (Fig. 7), we

find that these basin-scale nitrate \(\delta^{15}\)N differences also persist into the deep sea. (Note that

the basin-scale model nitrate \(\delta^{15}\)N differences shown in Fig. 7 are even larger for the nitrate

\(\delta^{15}\)N observations because the measurements are spatially biased towards areas of water

column denitrification in the Pacific and Indian Oceans (see Fig. 2).)

The remineralization of organic matter is one process that can—and has been used to—

explain both the elevation of deep Pacific nitrate \(\delta^{15}\)N (Peters et al., 2017; Rafter et al.,

2013; Sigman et al., 2009) and lowering of deep Atlantic nitrate \(\delta^{15}\)N (Knapp et al., 2008;

Marconi, 2017; Marconi et al., 2017; Tuerena et al., 2015) relative to the deep ocean mean.

Here we provide two additional pieces of evidence that argue for the remineralization of

organic matter as the key driver of these deep-sea nitrate \(\delta^{15}\)N differences. Our first piece

evidence is that the average subsurface source of nitrate to the Pacific and Indian Ocean

surface has a significantly higher \(\delta^{15}\)N (by 2‰ at the 250 m depth surface) than the

Atlantic and Southern Oceans (Figs. 5B and 7). Nitrate \(\delta^{15}\)N at 250 m is an admittedly

imprecise estimate for the nitrate upwelled to the surface, but even a slight elevation in

Pacific source nitrate \(\delta^{15}\)N and near complete nitrate utilization at the surface will translate

into higher sinking organic matter \(\delta^{15}\)N (i.e., see Fig. 1B).

Our second piece of evidence that the export and remineralization of organic matter drives

the inter-basin differences in deep nitrate \(\delta^{15}\)N comes from sediment trap measurements.

Averaging published sediment trap organic matter \(\delta^{15}\)N from the subtropical and tropical

Pacific gives a value of 8.5±2.9‰ (Knapp et al., 2016; Robinson et al., 2012), which is
significantly higher than measured in traps from the Atlantic (4.5±1.5‰) (Freudenthal et al., 2001; Holmes et al., 2002; Lavik, 2000; Thunell et al., 2004). Given observed Southern Ocean nitrate characteristics (Rafter et al., 2013), we estimate an even lower typical sinking organic matter δ15N of +1.5‰ for this region. (This assumes initial values are the Upper Circumpolar Deep Water and final values from the Open Antarctic Zone. Published results from the iron-fertilized Kerguelen Plateau region are predictably higher (Trull et al., 2008).) The much lower Southern Ocean sinking organic matter δ15N is consistent with partial consumption of nitrate at the surface (see Fig. 1B) and the entrainment of this nitrate in equatorward-moving intermediate waters acts to export nitrate with elevated δ15N to intermediate waters throughout the Southern Hemisphere (see discussion above).

Based on this evidence, it appears that global patterns of sinking organic matter δ15N are consistent with the remineralization of this organic matter driving subtle, but significant differences in deep-sea nitrate δ15N.

An alternative explanation for the deep-sea nitrate δ15N differences in Fig. 7 is that they reflect the lateral (along isopycnal) advection of elevated nitrate δ15N from ODZ regions. However, we can easily dismiss this explanation by looking at the meridional trends in deep-sea nitrate δ15N—following the deep waters from their entrance in the south and movement northward. What we find is that deep EANN nitrate δ15N (numbers in Fig. 5D) is lowest in the Southern Ocean and increases equatorward in the Pacific. Average observed nitrate δ15N below 2500 m increases from 4.7±0.1‰ in the Pacific sector of the Southern Ocean to 4.9±0.2‰ in the deep South Pacific, 5.4±0.2‰ in the deep tropical Pacific, and 5.2±0.2‰ in the deep North Pacific. This contrasts with no significant change in deep Atlantic nitrate δ15N, despite the export of slightly elevated nitrate δ15N into intermediate-depth Atlantic (see above and (Tuerena et al., 2015)) and the introduction of a different deep water mass (North Atlantic Deep Water) in the North Atlantic. The distribution of deep Pacific nitrate δ15N is coherent with elevated organic matter δ15N being produced and exported from the lower latitude surface and remineralized at depth. In other words, inter-basin differences sinking organic matter δ15N best explains the inter-basin differences in deep EANN and observed nitrate δ15N. Vertical, diapycnal mixing from the low latitude Pacific ODZ regions may also play a role in the south-to-north elevation of deep Pacific nitrate δ15N, but we cannot quantify the magnitude of that influence without a circulation model. Future work should look into this issue.

5 Conclusions
We find that an Ensemble of Artificial Neural Networks (EANN) can be trained on climatological distributions of physical and biogeochemical tracers to reproduce a global database of nitrate δ15N observations (Fig. 2) with good fidelity (Fig. 3). We used the EANN to produce global climatological maps of nitrate δ15N at a 1 degree-resolution from the surface to the seafloor. These results help identify spatial patterns (Figs. 4-6) and quantify regional and basin-average oceanic values of nitrate δ15N (Fig. 7). Major differences between the observed and EANN-predicted nitrate δ15N appear to be caused by temporal variability of nitrate δ15N in the upper ocean and in ODZs associated with variable nitrate uptake and denitrification rates. Additional measurements of nitrate δ15N will help to
develop seasonally-resolved maps that can improve upon the climatological mean map provided here.

Acknowledgments: M. Altabet, K. Casciotti, A. Santoro, B. Pasquier, as well as J. Granger and D. M. Sigman for (at-the-time) unpublished data. A complete list of references can be found in the Appendix. The compiled data set and data product is available in several online databases (BCO-DMO.org, pangaea.de, and webodv.awi.de). Color palettes were custom-made and are available via www.prafter.com.
References


Liu, K. K.: Geochemistry of inorganic nitrogen compounds in two marine environments: The Santa Barbara Basin and the ocean off Peru, University of Southern California, Los Angeles, 1979.


Figure 1: (A) A comparison of influences on average deep-sea nitrate (circle; concentration and δ^{15}N estimated here by EANN results in this work) including: the isotope effects of assimilation (yellow arrow), water column and sedimentary denitrification (red and gray arrows), and the addition of nitrate via remineralization of organic matter with higher and lower δ^{15}N (brown and purple arrows) (modified from Galbraith et al., (2008)). (B) An example of N isotopic fractionation on nitrate and organic matter biomass during nitrate assimilation in eastern equatorial Pacific surface waters (from Rafter and Sigman, (2016)). These calculations are based on isotopic fractionation equations of (Mariotti et al., 1981) simplified in (Sigman and Casciotti, 2001) with an isotope effect (ε) as shown in (A). The variable “f” is the observed / initial nitrate concentration.
Figure 2: The location of global nitrate $\delta^{15}$N observations used to constrain the Ensemble of Artificial Neural Networks are shown as red circles. Observations used as an 'external validation dataset' (those withheld from training the EANN) are shown in blue. Inset figure shows the number of observations versus depth. Contours are surface nitrate concentrations for October-December from World Ocean Atlas (Garcia et al., 2014).
Figure 3: The binned observed versus EANN-predicted nitrate $\delta^{15}$N are shown for all depths in (A) and for >1000 m in (B). The colors indicate the number of comparisons on the World Ocean Atlas grid. The anomalously high observed nitrate $\delta^{15}$N values (>30‰) in (A) are exclusively from the Eastern Tropical South Pacific waters (Bourbonnais et al., 2015; Casciotti et al., 2013; Rafter et al., 2012; Ryabenko et al., 2012). The difference (or residual) between the observations and EANN nitrate $\delta^{15}$N is made for all depths and the upper 1000 m in (C) with colors representing the dissolved oxygen content. Note the largest offsets between 100-500 m in (C) are associated with lowest oxygen content. Similarly, the mean residual (black) and Root Mean Square Error (RMSE; red) with depth (D) are highest in the upper 500 m. Dashed lines in (D) demonstrate the same statistics, but for the external validation dataset (blue in Fig. 2).
A Pacific Ocean nitrate $\delta^{15}$N [per mil]

$R^2 = 0.63$
RMSE = 2.55
mean bias = +0.02
n = 3092

$R^2 = 0.75$
RMSE = 0.25
mean bias = +0.08
n = 939
Figure 4: Zonal average of model nitrate $\delta^{15}N$ versus latitude or longitude for the: (A) Pacific Ocean, (B) Atlantic Ocean, (C) Southern Ocean, and the (D) Indian Ocean. White bars indicate no data because of land. The $R^2$, RMSE, mean bias, and total number (n) of observed versus EANN nitrate $\delta^{15}N$ are shown on the right for each region and depth range. White indicates regions of no data coverage.
Figure 5: (Left) Map view of nitrate $\delta^{15}$N from our EANN and our observations (circles) for the (A) average over the 0-50 m depth as well as the (B) 250 m, (C) 700 m, and (D) 3000 depth surfaces. Numbers are average model nitrate $\delta^{15}$N for that basin, depth, and latitude. (Right) Map views of nitrate $\delta^{15}$N error from the EANN model nitrate $\delta^{15}$N for the same depth surfaces on left.
Figure 6: Difference between the average nitrate $\delta^{15}\text{N}$ in the upper 50 m and 250 m depths in Figure 5. Dashed contours in low latitude ODZ regions and subtropical gyres indicate regions where nitrate $\delta^{15}\text{N}$ at 250 m is greater than the upper 50 m nitrate $\delta^{15}\text{N}$. 
Figure 7: Mean EANN nitrate $\delta^{15}$N (solid line) and 1-sigma standard deviation (envelope) with depth for each ocean basin. Note change in vertical and horizontal axes between top and bottom.
## Appendix: References for this version of seawater nitrate $\delta^{15}$N compilation

<table>
<thead>
<tr>
<th>Region</th>
<th>Year of Sampling</th>
<th>Month of Sampling</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacific North – Subarctic</td>
<td>Unknown sampling date</td>
<td>na</td>
<td>(Altabet and Francois, 1994)</td>
</tr>
<tr>
<td>Indian – Arabian Sea</td>
<td>1995</td>
<td>8</td>
<td>(Altabet et al., 1999)</td>
</tr>
<tr>
<td>Pacific North – Gulf of California</td>
<td>1990</td>
<td>6</td>
<td>(Altabet, 1999)</td>
</tr>
<tr>
<td>Pacific North – Subarctic</td>
<td>1971</td>
<td>7</td>
<td>(Wada, 1980)</td>
</tr>
<tr>
<td>Indian – Arabian Sea</td>
<td>1994</td>
<td>4</td>
<td>(Brandes et al., 1998)</td>
</tr>
<tr>
<td>Pacific North – ETNP</td>
<td>1993</td>
<td>12</td>
<td>(Brandes et al., 1998)</td>
</tr>
<tr>
<td>Pacific North – Tropical</td>
<td>1997</td>
<td>10 &amp; 11</td>
<td>(Voss et al., 2001)</td>
</tr>
<tr>
<td>Pacific North – Subarctic</td>
<td>2003</td>
<td>2</td>
<td>(Galbraith, 2007)</td>
</tr>
<tr>
<td>Atlantic North</td>
<td>2004</td>
<td>5</td>
<td>(Bourbonnais et al., 2009)</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>1996</td>
<td>5</td>
<td>(Sachs, 1999)</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>1998</td>
<td>1</td>
<td>(Pantoja et al., 2002)</td>
</tr>
<tr>
<td>Pacific North – Subarctic</td>
<td>2002</td>
<td>6</td>
<td>(Lehmann et al., 2005)</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>1977</td>
<td>6</td>
<td>(Liu, 1979)</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>2002 &amp; 2004</td>
<td>4 &amp; 5</td>
<td>(De Pol-Holz et al., 2009)</td>
</tr>
<tr>
<td>Pacific Tropical</td>
<td>2006</td>
<td>6</td>
<td>(Kienast et al., 2008)</td>
</tr>
<tr>
<td>Southern Ocean – Indian</td>
<td>2005</td>
<td>1 &amp; 2</td>
<td>(Trull et al., 2008)</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>2008 &amp; 2009</td>
<td>1, 2, &amp; 12</td>
<td>(Ryabenko et al., 2012)</td>
</tr>
<tr>
<td>Indian South</td>
<td>2011</td>
<td>10 &amp; 11</td>
<td>(Dehairs et al., 2015)</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>2012</td>
<td>11</td>
<td>(Bourbonnais et al., 2015)</td>
</tr>
<tr>
<td>Indian North</td>
<td>2007</td>
<td>9</td>
<td>(Gaye et al., 2013)</td>
</tr>
<tr>
<td>Atlantic South</td>
<td>2010 &amp; 2012</td>
<td>10 &amp; 1</td>
<td>(Tuarena et al., 2015)</td>
</tr>
<tr>
<td>Pacific South</td>
<td>2009</td>
<td>6</td>
<td>(Yoshikawa et al., 2015)</td>
</tr>
<tr>
<td>Region</td>
<td>Year(s)</td>
<td>Sections</td>
<td>References</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------</td>
<td>----------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Pacific North – SCS</td>
<td>1997</td>
<td>4</td>
<td>(Wong et al., 2002)</td>
</tr>
<tr>
<td>Arctic – Beaufort</td>
<td>2009</td>
<td>9</td>
<td>Granger unpublished</td>
</tr>
<tr>
<td>Atlantic North</td>
<td>2010</td>
<td>10 &amp; 11</td>
<td>Jenkins et al. Unpublished GEOTRACES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Knorr_199_leg4.pdf)</td>
</tr>
<tr>
<td>Atlantic Tropical</td>
<td>2010</td>
<td>2 &amp; 3</td>
<td>Frank et al. Unpublished GEOTRACES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(meteor81_1.pdf)</td>
</tr>
<tr>
<td>Pacific Tropical</td>
<td>2013</td>
<td>5 &amp; 6</td>
<td>(Lehmann et al., 2018)</td>
</tr>
<tr>
<td>Pacific North</td>
<td>2008</td>
<td>7</td>
<td>Granger Unpublished</td>
</tr>
<tr>
<td>Pacific North</td>
<td>2009 &amp; 2011</td>
<td>2 &amp; 7</td>
<td>(Umezawa et al., 2014)</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>2010 &amp; 2011</td>
<td>2-4</td>
<td>(Knapp et al., 2016)</td>
</tr>
<tr>
<td>Atlantic North – Subarctic</td>
<td>1989</td>
<td>6</td>
<td>(Voss, 1991)</td>
</tr>
<tr>
<td>Southern Ocean – Pacific</td>
<td>1995</td>
<td>4</td>
<td>(Sigman et al., 1999)</td>
</tr>
<tr>
<td>Southern Ocean – Indian</td>
<td>1995</td>
<td>1</td>
<td>(Sigman et al., 1999)</td>
</tr>
<tr>
<td>Southern Ocean – Pacific</td>
<td>2016</td>
<td></td>
<td>(Kemeny et al., 2016)</td>
</tr>
<tr>
<td>Pacific North – Tropical</td>
<td>2003</td>
<td>10</td>
<td>(Sigman et al., 2005)</td>
</tr>
<tr>
<td>Pacific North – ALOHA</td>
<td>2000</td>
<td>11</td>
<td>(Sigman et al., 2009)</td>
</tr>
<tr>
<td>Atlantic – North</td>
<td>2001-2002</td>
<td>1-12</td>
<td>(Knapp et al., 2005)</td>
</tr>
<tr>
<td>Atlantic – North</td>
<td>2002</td>
<td>10</td>
<td>(Knapp et al., 2008)</td>
</tr>
<tr>
<td>Pacific – North</td>
<td>2003</td>
<td>7 &amp; 8</td>
<td>(Knapp et al., 2011)</td>
</tr>
<tr>
<td>Indian - South</td>
<td>1999</td>
<td>1 &amp; 2</td>
<td>(Karsh et al., 2003)</td>
</tr>
<tr>
<td>Southern – Indian</td>
<td>1998 &amp; 1999</td>
<td>2,3,4,9, &amp; 12</td>
<td>(DiFiore et al., 2006)</td>
</tr>
<tr>
<td>Southern – Atlantic</td>
<td>2012</td>
<td>7</td>
<td>(Smart et al., 2015)</td>
</tr>
<tr>
<td>Atlantic – North</td>
<td>2011</td>
<td>10 &amp; 11</td>
<td>(Marconi et al., 2015)</td>
</tr>
<tr>
<td>Pacific – North</td>
<td>2004</td>
<td>7</td>
<td>(Casciotti et al., 2008)</td>
</tr>
<tr>
<td>Pacific – South</td>
<td>2005</td>
<td>11</td>
<td>(Casciotti et al., 2013)</td>
</tr>
<tr>
<td>Tropical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific – North</td>
<td>2003</td>
<td>11</td>
<td>(Casciotti and McIlvin, 2007)</td>
</tr>
<tr>
<td>Tropical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indian – Arabian</td>
<td>2007</td>
<td>9</td>
<td>(Martin and Casciotti, 2017)</td>
</tr>
<tr>
<td>Pacific – Tropical</td>
<td>2004-2007</td>
<td>3-12</td>
<td>(Rafter et al., 2012; Rafter and Sigman, 2016)</td>
</tr>
<tr>
<td>Region</td>
<td>Year</td>
<td>Pages</td>
<td>Source</td>
</tr>
<tr>
<td>------------------------</td>
<td>------</td>
<td>-------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Indian – Arabian Sea</td>
<td>2007</td>
<td>10</td>
<td>(DeVries et al., 2012) or Rafter and Sigman Unpublished</td>
</tr>
<tr>
<td>Pacific South – Tasman Sea</td>
<td>2010</td>
<td>1 &amp; 2</td>
<td>Rafter and Sigman Unpublished</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>2010</td>
<td>2</td>
<td>(Rafter et al., 2012)</td>
</tr>
<tr>
<td>Atlantic North – Subarctic</td>
<td>2010</td>
<td>4</td>
<td>Rafter and Sigman Unpublished</td>
</tr>
<tr>
<td>Pacific South</td>
<td>2005</td>
<td>1</td>
<td>(Rafter et al., 2013)</td>
</tr>
<tr>
<td>Pacific South – Tropical</td>
<td>2013</td>
<td>10-12</td>
<td>(Peters et al., 2017)</td>
</tr>
<tr>
<td>Atlantic North – Subarctic</td>
<td>2013</td>
<td>8</td>
<td>(Marconi et al., 2017)</td>
</tr>
<tr>
<td>Pacific North – Subarctic</td>
<td>1993</td>
<td>5</td>
<td>(Wu et al., 1997)</td>
</tr>
<tr>
<td>Arctic</td>
<td>2014</td>
<td>7 &amp; 8</td>
<td>(Fripiat et al., 2018)</td>
</tr>
</tbody>
</table>