Paleoceanography and Paleoclimatology

RESEARCH ARTICLE

10.1029/2023PA004612

Machine Learning Solutions to Regional Surface Ocean δ^{18}O-Salinity Relationships for Paleoclimatic Reconstruction

N. K. Murray[1], **A. R. Muñoz**[2], **and J. L. Conroy**[1,3]

[1]Department of Earth Science and Environmental Change, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA, [2]Department of Physics, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA, [3]Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA

Key Points:

- Mixed-layer regions with distinct salinity-seawater δ^{18}O paired values are identified by self-organizing map
- Inclusion of atmospheric or oceanic tracers in the self-organizing map shifts the spatial distribution of salinity-seawater δ^{18}O patterns
- Region identification is still dependent on available observational data, but the throughput offered here enables updating as data increase

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

N. K. Murray,
nkmurra2@illinois.edu

Citation:

Murray, N. K., Muñoz, A. R., & Conroy, J. L. (2023). Machine learning solutions to regional surface ocean δ^{18}O-salinity relationships for paleoclimatic reconstruction. *Paleoceanography and Paleoclimatology*, *38*, e2023PA004612. https://doi.org/10.1029/2023PA004612

Received 18 JAN 2023
Accepted 14 AUG 2023

Abstract Stable isotope-based reconstructions of past ocean salinity and hydroclimate depend on accurate, regionally constrained relationships between the stable oxygen isotopic composition of seawater ($\delta^{18}O_{sw}$) and salinity in the surface ocean. An increasing number of $\delta^{18}O_{sw}$ observations suggest greater spatial variability in this relationship than previously considered, highlighting the need to reassess these relationships on a global scale. Here, we use available, paired $\delta^{18}O_{sw}$ and salinity data ($N = 11{,}119$) to create global interpolations of each variable. We then use a self-organizing map, a specialized form of machine learning, to define 19 regions with unique $\delta^{18}O_{sw}$-salinity relationships in the surface (<50 m) ocean. Inclusion of atmospheric moisture-related variables and oceanic tracer data in additional self-organizing map experiments indicates global surface $\delta^{18}O_{sw}$-salinity spatial patterns are strongly forced by the atmosphere, as the SOM spatial output is highly similar despite no overlapping input data. Our approach is a useful update to the previously delimited regions, and highlights the utility of neural network pattern extraction in spatiotemporally sparse data sets.

Plain Language Summary Our understanding of past changes in the ocean and atmosphere often comes from information stored in biological marine carbonate archives. Stable oxygen isotope measurements from such archives record past ocean temperature as well as the stable oxygen isotopic composition of seawater. In the surface ocean, stable oxygen isotope values have a strong, linear, relationship with salinity, thought to be the result of evaporation and precipitation affecting stable isotope ratios and salinity in a similar manner. This permits reconstruction of past ocean salinity from marine carbonates, which may be used to infer large-scale hydroclimate variability, in some cases. However, this manner of reconstruction relies on the linear isotope-salinity relationship, which remains poorly constrained and is influenced by a combination of oceanic and atmospheric processes. In this work, we define regional isotope-salinity relationships and broadly assess whether ocean or atmospheric processes play a larger role in the spatial distribution of these regions. We identify new regions compared to previous work, and find that over most of the ocean, precipitation, precipitation isotope values, and evaporation are important in setting seawater isotope-salinity patterns.

1. Introduction

Stable oxygen isotope (δ^{18}O) values of fossil marine carbonates are widely used to understand past climate and ocean variability, including changes in sea surface temperature and salinity on interannual to million-year timescales (e.g., Abram et al., 2020; Chen et al., 2018; Gorman et al., 2012; Grothe et al., 2020; Holbourn et al., 2021; Reed et al., 2022). These δ^{18}O values are primarily a function of ocean temperature and the stable oxygen isotope composition of seawater ($\delta^{18}O_{sw}$) during formation (Epstein et al., 1953), with a smaller contribution from seawater pH (Krief et al., 2010; Robbins et al., 2017; Spiro et al., 1997). As the temperature contribution can be assessed independently through paired measurements of δ^{18}O and Sr/Ca in corals or δ^{18}O and Mg/Ca in foraminifera (Alibert & McCulloch, 1997; Gagan, 1998; Mitsuguchi et al., 1996), $\delta^{18}O_{sw}$ can be estimated, given that the temperature-dependent fractionation between seawater and biogenic carbonate ($\delta^{18}O_{sw}$-$\delta^{18}O_{c}$) is well-defined (e.g., Epstein et al., 1953; O'Neil et al., 1969). The resulting $\delta^{18}O_{sw}$ reconstructions can provide valuable information on past salinity, as $\delta^{18}O_{sw}$ and salinity are strongly related in the surface mixed layer (Craig & Gordon, 1965). In turn, salinity may provide information on atmospheric moisture balance (Durack et al., 2012), as well as ocean circulation (Yu et al., 2020).

The linear $\delta^{18}O_{sw}$-salinity relationship, observed across different water masses and regions, varies due to ocean circulation, water mass mixing, and disparate regional hydrologic processes. Regional atmospheric moisture balance, or evaporation minus precipitation (E-P), is a key control on sea surface salinity, and thus $\delta^{18}O_{sw,}$ as these processes impact both variables in a similar manner. However, the additional sensitivity of $\delta^{18}O_{sw}$ to precipitation and runoff with varying $\delta^{18}O$ values can alter the slope and intercept of the $\delta^{18}O_{sw}$-salinity relationship. The $\delta^{18}O$ signature of regional precipitation ($\delta^{18}O_p$) and runoff is thus integrated into the $\delta^{18}O_{sw}$ values of the mixed layer. This influx reflects many influences, such as the seasonality of precipitation $\delta^{18}O$ values, moisture source regions, rain reevaporation, cloud type, and cloud microphysics, none of which influence salinity (Konecky et al., 2019). Additionally, ocean advection and upwelling can further influence the mixed layer $\delta^{18}O_{sw}$-salinity relationship (Benway & Mix, 2004; Conroy et al., 2023; Stevenson et al., 2018). In high latitudes, seasonal sea ice melt/formation and runoff can also affect the $\delta^{18}O_{sw}$-salinity relationship (Bauch et al., 2005; Dubinina et al., 2017, 2019). The *y*-intercept of the $\delta^{18}O_{sw}$-salinity relationship, the "freshwater endmember," is thought to be an integrated measure of evaporation, precipitation, and the isotope signal of each, providing additional information about regional hydroclimate (Delaygue et al., 2001), along with the slope, which is critical for isotope-based proxy forward models of marine paleoclimate archives (Thompson et al., 2022). Consequently, it is essential to have an accurate understanding of the modern spatial variability in $\delta^{18}O_{sw}$-salinity relationship to begin to tease apart its numerous controls and to provide a framework suitable for stable isotope-based paleoceanographic and paleoclimate interpretations.

Over the last two decades, the NASA GISS Seawater Oxygen Isotope database has been the primary source for $\delta^{18}O_{sw}$ data (Schmidt, 1999). These data were used to define widely used regional mixed layer, intermediate, and deep water mass $\delta^{18}O_{sw}$-salinity relationships (LeGrande & Schmidt, 2006). For the mixed-layer regions, surface circulation patterns and regions "expected to have distinct $\delta^{18}O$ signatures" were used to geographically partition the data to define $\delta^{18}O_{sw}$-salinity relationships. In this sense, geographical and ocean circulation features were used to infer the spatial boundaries of regions with specific $\delta^{18}O_{sw}$-salinity relationships. Sixteen mixed layer regions were calculated from linear regressions of the upper 50 m of observational data, though the boundaries of these regions were not provided (LeGrande & Schmidt, 2006). Use of these regions is further complicated by recent model and observational data indicating additional spatial variability not captured by these mixed-layer regions (e.g., Bauch et al., 2005; Conroy et al., 2017; Dubinina et al., 2017; Russon et al., 2013; Stevenson et al., 2015). This is partly due to the limited availability of global, paired $\delta^{18}O_{sw}$-salinity data, but may also be related to the method of region identification. Ultimately, this limited understanding of $\delta^{18}O_{sw}$-salinity relationships inhibits accurate regional interpretations of marine carbonate $\delta^{18}O$ records.

Here, we build on previous work by using a specialized form of neural network, the self-organizing map (SOM), to define regions with specific, paired $\delta^{18}O_{sw}$-salinity values (Akman et al., 2019; Kohonen, 1982; Liu et al., 2006). The SOM approach can be particularly useful in the earth sciences, as it can extract complex and nonlinear patterns in time or space from sparse, multidimensional data sets (Liu et al., 2006; Liu & Weisberg, 2011). In this work, we provide 1° × 1° gridded, global $\delta^{18}O_{sw}$ and salinity data sets from paired $\delta^{18}O_{sw}$ and salinity observations to a feed-forward back propagation neural network. We then use the SOM, as implemented in the SOMPY library, as a tool for region identification based on spatial patterns in the model output (Moosavi et al., 2014). We apply contour mapping on the SOM output to define regional spatial boundaries of 19 regions and we then assess the $\delta^{18}O_{sw}$-salinity relationships in these regions. Finally, we evaluate how the spatial patterns for regional $\delta^{18}O_{sw}$-salinity relationships change when we include key atmospheric and oceanic tracer data along with $\delta^{18}O_{sw}$-salinity data in additional SOM experiments.

2. Materials and Methods

2.1. Data Sources and Parameters

Data for this project come from three publicly available $\delta^{18}O_{sw}$ databases. We utilize data from the NASA GISS Seawater Oxygen-18 Database, GEOTRACES, and L'OCEAN, along with recent data from the tropical Pacific (Anderson et al., 2014; Conroy et al., 2017; Reverdin et al., 2022; Schmidt et al., 1999). We select samples with only paired $\delta^{18}O_{sw}$ and salinity measurements ($N = 11{,}119$) for the 1° × 1° gridded neural network interpolations (Figure 1). This maintains direct spatial and temporal overlap of the input observational data and provides the same "starting point" for the interpolation. Data are filtered by depth (<50 m) to approximate the mixed layer of the ocean and we set a threshold of 20 PSU to remove strongly freshwater-influenced datapoints. We chose the estimate of mixed layer depth in line with the methods of LeGrande and Schmidt (2006). However, we refer to our data as the "surface ocean," to acknowledge that a boundary set at 50 m does not perfectly capture regional

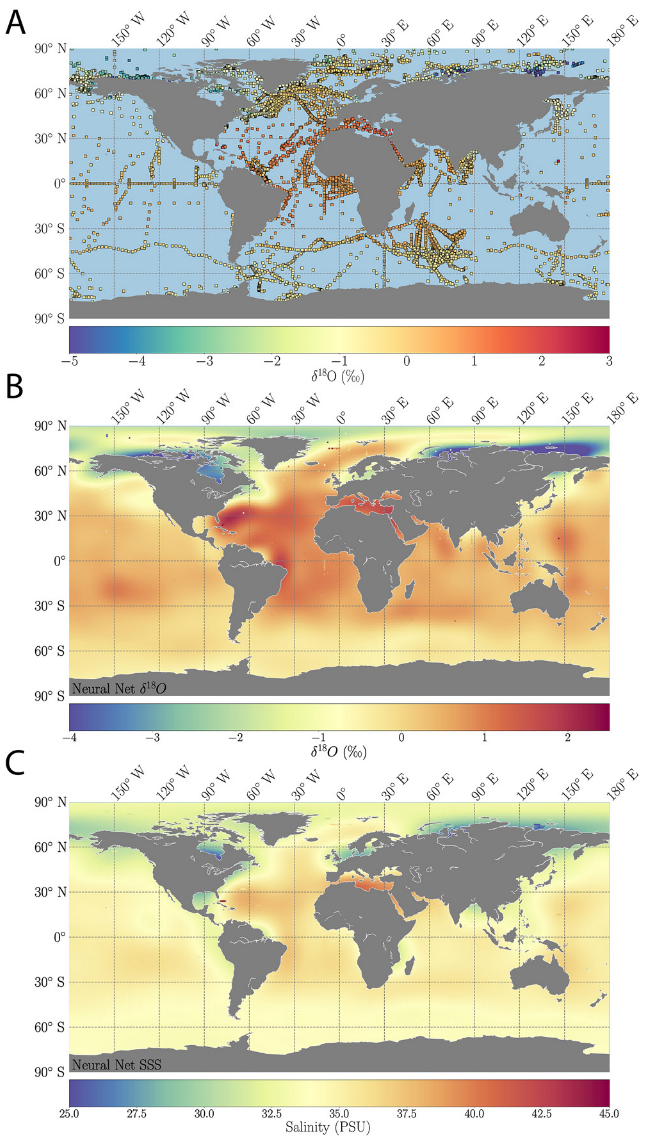

Figure 1. (a) Map of paired $\delta^{18}O_{sw}$-salinity datapoints used in this study. Color bar denotes $\delta^{18}O_{sw}$ value (‰, VSMOW) for each point. $1^\circ \times 1^\circ$ feed-forward, back propagation interpolations of (b) $\delta^{18}O_{sw}$ and (c) salinity. Spatial patterns are similar globally, as is expected from the linear relationship between $\delta^{18}O_{sw}$ and salinity, but with notable regional differences.

or seasonal changes in mixed layer depth. Temporally, the $\delta^{18}O_{sw}$ and salinity data set spans 1950–2021. Due to the sparsity of the observations, all data are treated as representative of the long-term mean, similar to previous work (LeGrande & Schmidt, 2006).

2.2. Neural Network

To build a $\delta^{18}O_{sw}$-salinity data set, we use a feed-forward back propagation neural network with two hidden layers to interpolate the two global data sets for $\delta^{18}O_{sw}$ and salinity (Figure 1 and Figure S1 in Supporting Information S1). We use 85% of the data for training and 15% for validation, the optimum ratio and the training precedent for machine learning (Bishop, 2006). During the training process, 85% of the data is given to the neural network, which iterates until the weights in the hidden layer have converged to produce a fit to the data. The weights are trained using AdaGrad, the Adaptive Gradient algorithm, a stochastic gradient descent method with a dynamic learning rate. This ensures consistent convergence and removes the need for tuning the learning rate hyperparameter (Duchi et al., 2011). After the training step, the remaining 15% of the data is incorporated into the input. Using the trained model, we feed 100% of the data back into the model to verify that the model is predictive beyond the training data. The training output and the validation output are compared to compute a total error for the neural network. This verification step demonstrated that our model had a 2.9% error on the training data, and a 3.8% error on the test data. The increase in the error indicates that the model is reasonably generalizable without having overfit to the training data. Plots showing the error of each point with respect to the neural network model are shown in Figures S2 and S3 in Supporting Information S1.

In the neural network used here, the input layer consists of four variables: $\delta^{18}O_{sw}$ or salinity, latitude, longitude, and sample depth. The first hidden layer is composed of 1,500 nodes, and the second layer is composed of 500 nodes. The architecture was converged by scanning over each layer depth in increments of 250 until the model error was reasonable and matched coastal features. The final output of each neural network provides a $1^\circ \times 1^\circ$ global data set for $\delta^{18}O_{sw}$ and salinity (Figure 1).

2.3. Self-Organizing Map

The SOM is a specialized form of a neural network which takes in data as a grid of nodes that spatially adjust during training. SOMs are used for dimensionality reduction and have been shown to identify patterns with noisy data and nonlinear trends better than empirical orthogonal functions (Liu et al., 2006). During data processing, the SOM uses competitive learning to take each data vector (the data at each $1^\circ \times 1^\circ$ grid point) from the interpolations and sort it into a "best matching unit" (BMU) based on the dimensions of the data (i.e., depth, $\delta^{18}O_{sw}$, and salinity). There is thus no numeric information in the BMUs, they only indicate similarity of data points as measured by a Euclidean distance in feature space. This provides information about the relationships within the data set, that is, what grids are similar to one another, based on the input data.

For our application, we are concerned with how the feature space (depth, $\delta^{18}O_{sw}$, and salinity) maps to a single classifying variable, the BMU. In our analysis, we assume that large changes in feature space will map to distinct spatial regions where the boundaries between regions are marked by large BMU changes over small spatial extents. This approximation allows us to identify surface ocean regions with large amounts of data that congregate in feature space but lacks the ability to distinguish distinct linear relationships that are very similar in feature space. For example, this method excels at identifying the large, subtropical North Pacific gyre region, but will

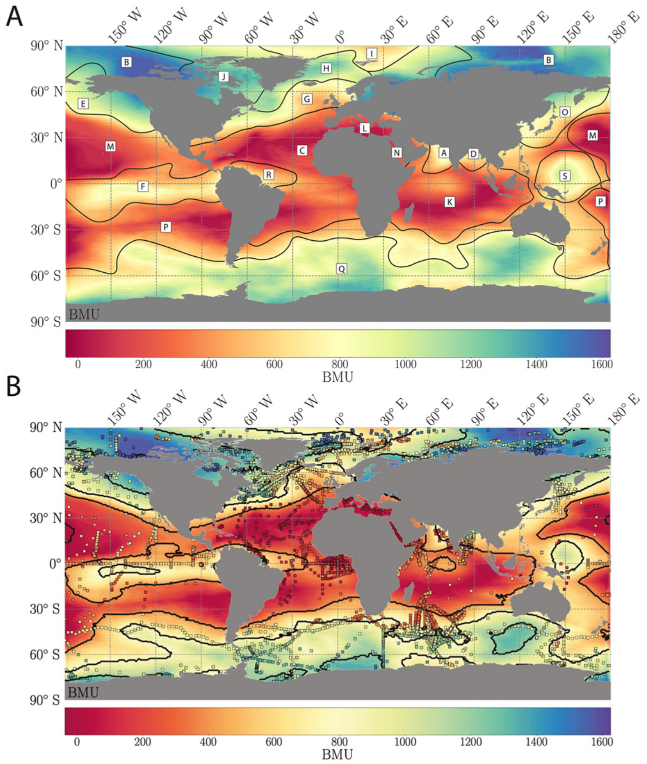

Figure 2. (a) Self-organizing map (SOM) output for the global ocean with regions corresponding to Table 1 labeled. The SOM is generated from 1° × 1° grid interpolations of $\delta^{18}O_{sw}$ and salinity and the boundaries between regions are taken from contour lines applied in (b). SOM output is in feature space and has no numeric value, however a scale showing the range of best matching units (BMU) is provided. Grid point color shows the BMU and relationships between grid points. Regions with similar colors have similar data in feature space: $\delta^{18}O_{sw}$, and salinity (e.g., red regions are similar to other red regions). (b) SOM validation output after observational data is fed back into the SOM model. Observational datapoints are shown as squares and the datapoint color indicates the BMU. Observational datapoint BMUs generally correspond with interpolation BMUs (background SOM output). This indicates that the $\delta^{18}O_{sw}$ and salinity interpolations are good models of the observational data.

not show small subsets of that data which may fall off the overall regional relationship, such as a single cruise in a single season in the North Pacific.

2.3.1. Self-Organizing Map: Training and Error

The SOM is trained in two parts: a coarse stage and a fine stage. During training, we designate an initialization mechanism, a number of nodes, a number of training iterations (both a fine and coarse stage) and calculate two error values: the quantization and topographic error. Both errors are unitless. The quantization error is the sum of the Euclidean distances between the data points and their respective BMUs. The topographic error evaluates what proportion of the data points do not have its first or second BMU as a neighboring unit. During training, the rough and fine training lengths were adjusted to reduce topographic error and quantization error. During this same step, the SOM was initialized by one of two methods: random or principal components. With a random initialization, it is often challenging to get reproducible output from the SOM (Schlegel et al., 2017). Accordingly, we use a principal component initialization, which uses a principal component analysis calculated from the provided data set and creates a reproducible output. The initial weights are normalized by the variance of the data to ensure numerical stability. We determine that a coarse training length of 30 iterations and a fine training length of 10 iterations converges the errors to a topographic error of 0.00 and a quantization error of 0.04. We train the SOM on our $\delta^{18}O_{sw}$ and salinity interpolation data sets. The result is a more robust SOM output than what could be generated from the 11,119 observational data points alone, as there is significantly more data for training. The SOM output from the interpolated data points and the output from the observed data points are in good agreement (Figure 2), indicating the $\delta^{18}O_{sw}$ and salinity interpolations are accurate representations of the observational data set.

Finally, a key decision in the training process is designating a number of nodes in the SOM, or the "grid size," which determines the number of BMUs. With too few nodes, the resolution is too coarse, potentially leading to missed regions. With too many nodes, there is a negligible effect, but there would be many regions with too few points to perform linear regressions. There is no risk of overfitting unless the number of nodes equals the number of data points or more. In this case, each 1° × 1° grid datapoint would be in its own BMU and there would be no regional relationship. We used 40 nodes (i.e., a 1,600 BMU grid) as this is the minimum number of nodes needed to converge the SOM. After 40 nodes, no change is apparent in the output, therefore no benefit is gained from additional nodes.

2.3.2. Self-Organizing Map: Region Identification

Regional boundaries coincide with BMU color gradients and are most reliable in the areas with large amounts of observational data (Figure 2). The SOM does not define sharp boundaries between regions with no observational data separating them. A salient example of a large interpolation distance with no observational data is the southeastern subtropical Pacific. Such locations with limited or no observational data are unable to be statistically identified as distinct from or similar to surrounding regions.

In addition to the initial SOM, based solely on $\delta^{18}O_{sw}$ and salinity, we run three additional SOM experiments that include ocean and atmospheric variables known to affect $\delta^{18}O_{sw}$ and salinity values. In one experiment, we also include an "ocean tracer" that includes nitrate, phosphate, silicate and dissolved oxygen circulation tracers sourced from the NOAA World Ocean Atlas, a suite of quality controlled global ocean data sets (Garcia et al., 2006). In the second experiment, we add an "atmospheric tracer" that includes global E-P reanalysis data from ERA-Interim, a global atmospheric reanalysis data set, and $\delta^{18}O_p$ (Dee et al., 2011; IAEA). For this SOM experiment, long-term annual mean $\delta^{18}O_p$ data was sourced from the International Atomic Energy Agency Global

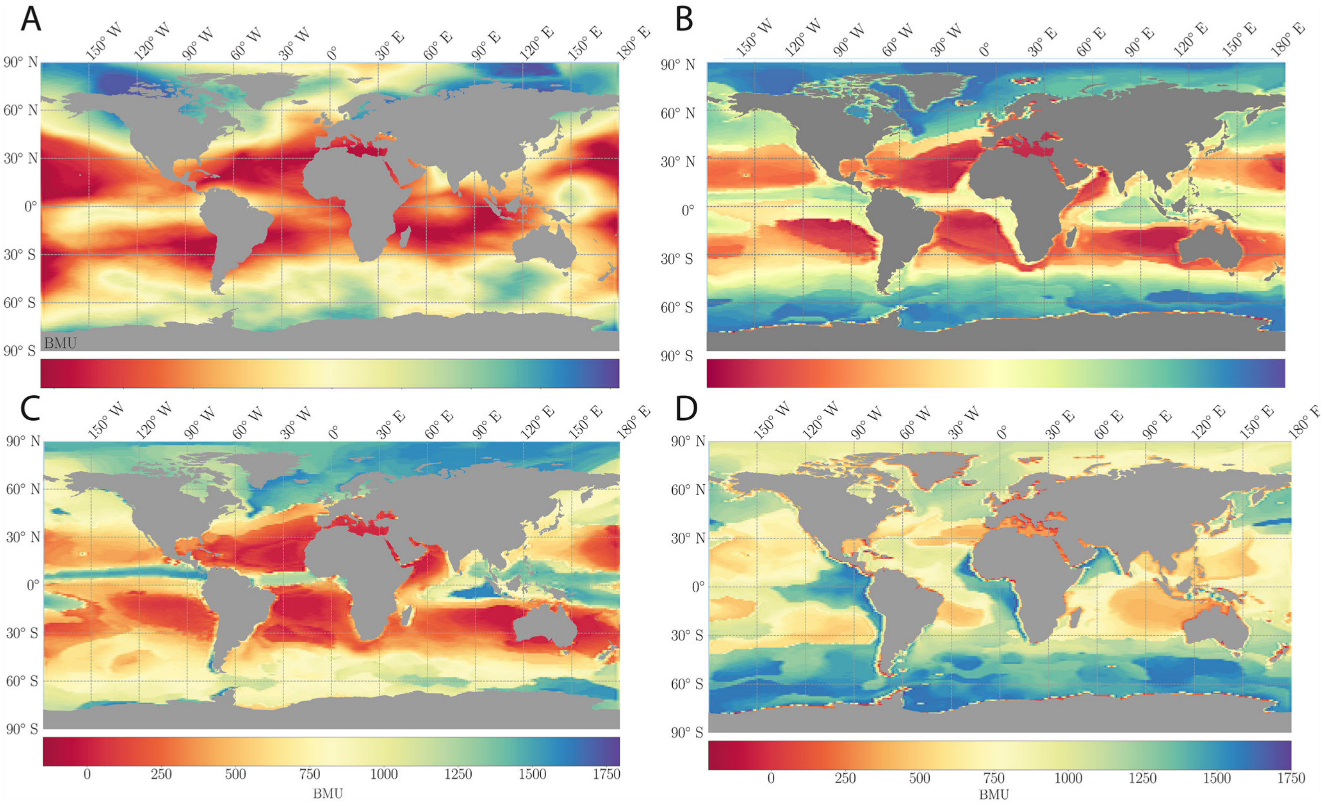

Figure 3. (a) SOM with $\delta^{18}O_{sw}$ and salinity data (same as in Figure 2a). (b) “All tracer” SOM generated with $\delta^{18}O_{sw}$, salinity, $\delta^{18}O_p$ (GNIP annual mean (1953–2021), all sites), E-P (long-term annual mean (1979–2016) mm/day from ERA-Interim reanalysis data set), and ocean tracers to 50 m (phosphate, nitrate, dissolved oxygen, and silicate, accessed from NOAA World Atlas 05). (c) “Atmospheric tracer” SOM run with only atmospheric variables: $\delta^{18}O_p$, E-P. (d) “Ocean tracer” SOM with only oceanic tracers (nitrate, phosphate, dissolved oxygen, and silicate).

Network of Isotopes in Precipitation (IAEA GNIP), and interpolated to 1° × 1° using PyKrige, a Python library used for kriging in geoscience data sets. We use a Gaussian variogram and parameters determined by an L1 norm minimization. Finally, an SOM containing the ocean tracers, atmospheric tracers and the original $\delta^{18}O_{sw}$ and salinity data (“all tracer” SOM) was created (Figure 3b).

2.4. Regional $\delta^{18}O_{sw}$-Salinity Fittings

We calculate linear regressions on the $\delta^{18}O_{sw}$ and salinity observational datapoints in each SOM-defined region (Table 1). The observational data used for each regional linear regression is decided by data points that fall into the contour-identified regions of the SOM. Contours were calculated using the “serial” algorithm from ContourPy. The number of contours was determined by using clearly separated regions (e.g., Mediterranean Sea, Southern Ocean) as a threshold. Spurious signals (i.e., contour “dots” enclosing areas smaller than 2° × 2° grid points) were removed for clarity and ease of use. Secondary regressions are also reported for a region when multiple relationships are apparent and $N \geq 10$. These secondary relationships were visually identified and typically are derived from single data sets that represent a smaller sub-region or time period. We do not consider values that do not represent open-ocean conditions (e.g., samples from lagoonal environments), although these points are still included in the figures.

3. Results

Broadly, all four SOM outputs (Figure 3) share core features, but with significant differences reflective of the input data sets. All SOM outputs reflect the subtropical gyres in low BMUs (red/orange), separations between them, and high latitude regions on the opposite end of the BMU spectrum (blue-green). Generally, the boundaries between regions are sharper in the SOM models with atmospheric or oceanic tracers (Figures 3b–3d) compared

Table 1
Regional $\delta^{18}O_{sw}$-Salinity Relationships Calculated in This Study

Map ID	Region name	Sample size (N)	Slope	Intercept	R^2	Slope error (1σ)	Intercept error (1σ)	RMSE (‰)
A	Arabian Sea Primary	302	0.17	−5.52	0.42	0.01	0.41	0.12
	Arabian Sea: Secondary	*17*	*0.57*	*−19.08*	*0.44*	*0.17*	*5.82*	*0.24*
B	Arctic Ocean	874	0.57	−19.94	0.80	0.01	0.29	0.98
C	Atlantic Ocean Primary	932	0.18	−5.57	0.54	0.01	0.20	0.16
	Atlantic Ocean Secondary: Delaware Coastal Current	*29*	*0.18*	*−6.87*	*0.88*	*0.01*	*0.40*	*0.14*
	Atlantic Ocean Secondary: Bahamas, Florida, Mid-Atlantic coast	*114*	*0.13*	*−2.89*	*0.56*	*0.01*	*0.36*	*0.45*
D	Bay of Bengal	112	0.19	−6.22	0.87	0.01	0.23	0.17
	Bay of Bengal Secondary	*16*	*0.24*	*−8.77*	*0.78*	*0.03*	*1.06*	*0.30*
E	Northeastern Pacific/Bering Sea/Chukchi Sea	608	0.37	−13.01	0.48	0.02	0.49	0.72
F	Central/Eastern Equatorial Pacific	720	0.14	−4.46	0.70	0.00	0.12	0.11
G	Northeast Atlantic	569	0.18	−5.91	0.91	0.00	0.08	0.16
H	GIN Seas, Northeast Atlantic	2,072	0.32	−11.07	0.61	0.01	0.19	0.5
I	Svalbard	197	0.19	−6.47	0.39	0.02	0.58	0.26
	Svalbard*: secondary*	*20*	*0.06*	*−1.97*	*0.09*	*0.05*	*1.59*	*0.26*
J	Hudson Bay, Hudson Strait, Baffin Bay, Northwestern Passages	622	0.44	−15.70	0.49	0.02	0.57	0.65
K	Indian Ocean	429	0.35	−11.78	0.43	0.02	0.69	0.11
	Indian Ocean Secondary	*25*	*0.12*	*−3.30*	*0.57*	0.02	0.68	0.28
L	Mediterranean Sea	147	0.226	−8.66	0.76	0.01	0.47	1.42
M	Tropical to Subtropical North Pacific Primary	188	0.55	−18.70	0.92	0.01	0.48	0.11
	Tropical to Subtropical North Pacific Secondary: SPURS2 cruise	45	0.03	−0.78	0.02	0.03	1.01	0.10
N	Red Sea	48	0.30	−10.27	0.95	0.01	0.38	0.10
O	Sea of Okhotsk	282	0.44	−15.00	0.94	0.01	0.21	0.24
P	South Pacific	36	0.33	−11.27	0.63	0.04	1.56	0.22
Q	Southern Ocean	1,942	0.40	−13.76	0.58	0.01	0.26	0.15
R	Western Tropical Atlantic	404	0.17	−5.25	0.76	0.00	0.16	0.44
S	Western Pacific	340	0.28	−9.28	0.76	0.01	0.29	0.16
	Western Pacific Secondary: New Guinea	*10*	*0.12*	*−4.22*	*0.39*	0.05	1.93	0.09

Note. Map ID corresponds to those labeled in Figure 2. Secondary and tertiary slopes are shown in italics.

to the output generated from the paired $\delta^{18}O_{sw}$-salinity alone (Figure 3a). The ocean tracer data set has a smaller range in BMU values and less spatial variability (Figure 3d). The Intertropical Convergence Zone (ITCZ) is only readily apparent in the SOM output with $\delta^{18}O_p$ and E-P data (Figure 3c), not in the $\delta^{18}O_{sw}$-salinity output alone (Figure 3a).

We use the "$\delta^{18}O_{sw}$-salinity only" SOM as the focus of our discussion on regional relationships to compare with previously defined regions and to avoid prior assumptions and biases in the SOM. Although we include the atmospheric and oceanic variables in our secondary SOM experiments to assess their contributions to $\delta^{18}O_{sw}$-salinity values, using these versions of the SOM for region delineation would include the prior assumption that atmospheric or oceanic variables are a significant control globally, which may not be an appropriate umbrella assumption. However, regional boundaries and relationships could certainly defined from the multitracer SOMs, for specific regional studies where the impact of the atmospheric or oceanic forcing is significant.

We identify 19 regions from the "$\delta^{18}O_{sw}$-salinity only" SOM (Figure 2 and Table 1). Several of these regions were not previously identified in LeGrande and Schmidt (2006), though some are identified in regionally focused studies, such as the western tropical Pacific and the Russian Arctic seas (Bauch et al., 2005; Conroy et al., 2017; Dubinina et al., 2017). Slope values range from −0.07‰ to 0.57‰/PSU and intercept (freshwater endmember)

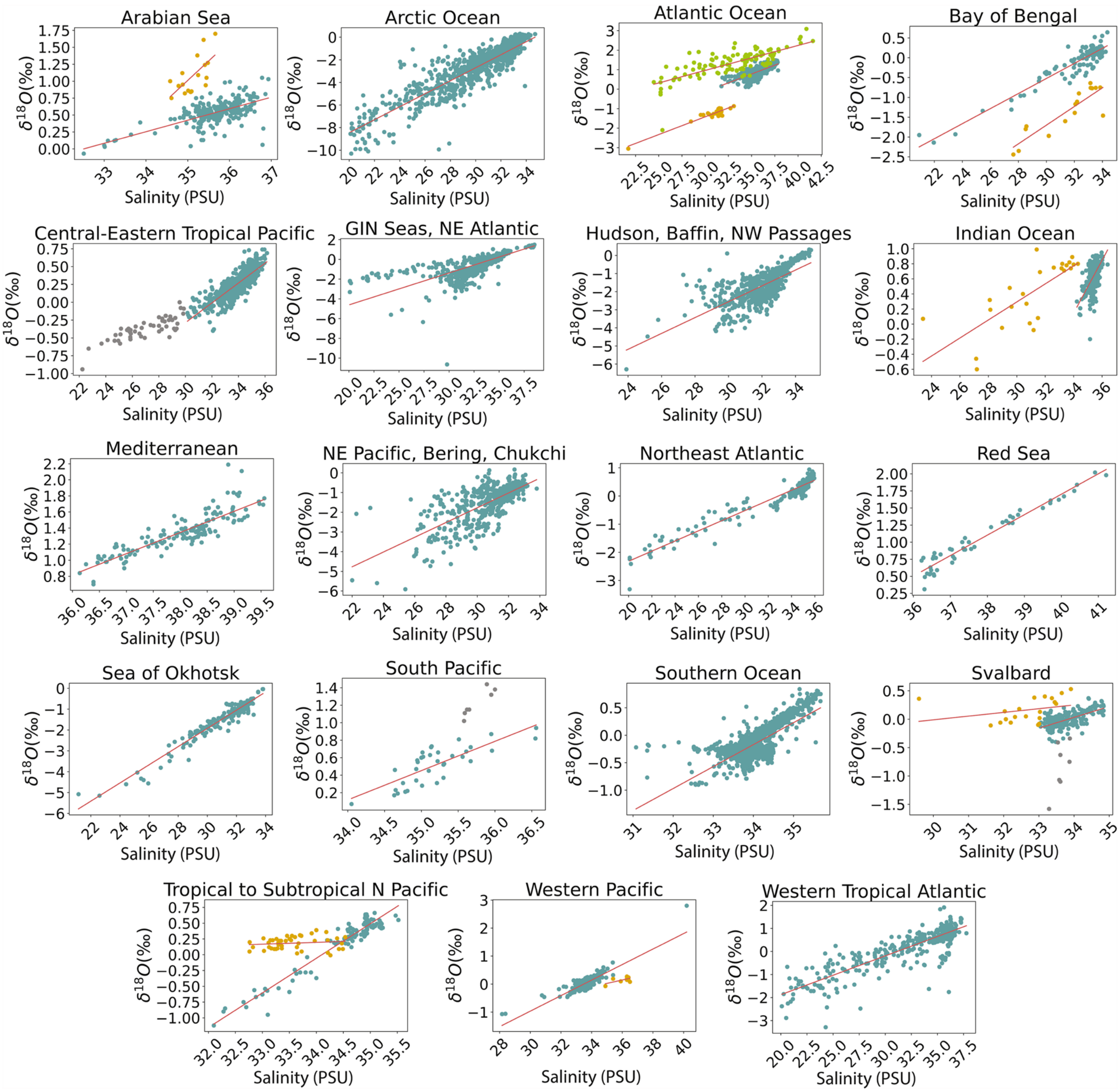

Figure 4. Linear regressions for regions in Table 1. Primary relationships are shown in green-blue. Gray points are considered outliers and not included in regressions; secondary regressions and associated values are shown in yellow and tertiary regressions are shown in light green.

values range from −19.95‰ to −0.79‰, with the slopes and freshwater endmembers significantly correlated ($r = -0.99$, i.e., low slopes coincide with more positive intercepts) (Figure 4). Generally, R^2 values are moderate to high and most relationships are statistically significant, with the exception of one high latitude relationship (Svalbard secondary) and one cruise (SPURS2) in the North Pacific (Table 1). Higher slopes tend to have higher R^2 values ($r = 0.35$).

4. Discussion

4.1. SOM Patterns and Regional $\delta^{18}O_{sw}$-Salinity Relationships

Four SOM outputs are produced to provide insight into the magnitude of influence between atmospheric versus oceanic drivers in the $\delta^{18}O_{sw}$-salinity relationship. To do this, we compared the "ocean tracer" and "atmospheric tracer" (Figures 3c and 3d) output to the "$\delta^{18}O_{sw}$ and salinity only" output (Figure 3a). The BMU range of all outputs is the same, but derived from different input data sets. Despite this, the spatial patterns in the "$\delta^{18}O_{sw}$ and salinity only" and "atmospheric tracer" output share many features and are notably similar, with the exception of the prominent equatorial band (the ITCZ) in the latter. This suggests that globally, the $\delta^{18}O_{sw}$-salinity relationship is highly influenced by $\delta^{18}O_p$ and E-P, which are the data sets used for the atmospheric tracer. The "ocean tracer" output, by contrast, is more dissimilar to the "$\delta^{18}O_{sw}$ and salinity only" output and the ocean tracer's inclusion in the "all tracer" output (Figure 3b) leads to more spatial differences from the "$\delta^{18}O_{sw}$ and salinity only" output than atmospheric alone. While both the tracers produced in this work are imperfect representations of all real-world ocean and atmospheric processes, our data imply that atmospheric processes strongly influence $\delta^{18}O_{sw}$ and salinity data globally. From the SOM tracer experiments, we move forward using the "$\delta^{18}O_{sw}$ and salinity only" output, henceforth called simply, "the SOM."

Generally, the SOM output partitions the global ocean into two endmembers, visualized by red-orange or blue-green BMUs (Figure 2). Although the SOM output is defined by both salinity and $\delta^{18}O_{sw}$ values, there is a broad pattern in the BMUs that is reflective of salinity, given the strong covariation of these variables. Here we focus on the differences, which are reflected in different regional slopes and freshwater endmembers (intercepts). Generally, low slopes and more positive freshwater endmembers are expected in the tropics due to moisture source proximity minimizing isotopic distillation and maintaining relatively higher $\delta^{18}O_p$ values. This leads to a smaller change in $\delta^{18}O_{sw}$ per unit change in salinity and more positive freshwater endmember values (LeGrande & Schmidt, 2006). Indeed, most tropical slopes are low (~0.2–0.3). However, $\delta^{18}O_{sw}$-salinity slopes in the tropics can also be influenced by specific, local processes. For example, the higher slope in the western Pacific relative to the eastern and central Pacific may reflect the impact of lower $\delta^{18}O_p$ values in this area with higher precipitation amounts, leading to a larger change in $\delta^{18}O_{sw}$ per unit salinity (Conroy et al., 2017).

High latitude regions have both the highest and lowest slopes in the data set (>0.5 and < 0.2). Sea ice melt and formation are implicated as key drivers of the $\delta^{18}O_{sw}$-salinity relationships in these areas (Bauch et al., 2005, 2010; Dubinina et al., 2017; LeGrande & Schmidt, 2006). A sea ice signal may manifest in two parts: during formation, we expect large increases in salinity associated with brine exclusion (Rohling, 2013). During melting, salinity decreases. In both cases, we expect low slopes or slopes near zero as stable oxygen isotope fractionation between sea ice and water is very small, but salinity changes remain comparatively large, as seawater and sea ice salinities are often >10 PSU different. Large modifications in salinity may even lead to a decoupling of the $\delta^{18}O_{sw}$-salinity correlation (Dubinina et al., 2017), where the range in salinity is so comparatively large compared to $\delta^{18}O_{sw}$ that there is no significant relationship. Low R^2 values in such high latitude regions may thus be due to sea ice related freshwater fluxes, leading to $\delta^{18}O_{sw}$-salinity decoupling. Regional studies of high latitude seas have also identified continental runoff from rivers with very low $\delta^{18}O$ values as an additional influence on $\delta^{18}O_{sw}$ (Bauch et al., 2005; Dubinina et al., 2017, 2019). Such low $\delta^{18}O$ values in high latitude runoff are a result of precipitation and streamflow with low $\delta^{18}O$ values (Nan et al., 2019). Thus, in the high latitude regions with high slopes and more negative intercepts, runoff is likely a major control on the $\delta^{18}O_{sw}$-salinity relationship.

Several regions defined from the SOM output contain secondary slopes that can be clearly separated from the primary slopes (Figure 4 and Table 1). Some secondary regression equations may represent different water masses (e.g., Akhoudas et al., 2021; Benway & Mix, 2004). However, based on the available metadata, spatial and temporal variability are the main cause of the SOM-region secondary slopes. For example, spatially, the Delaware Coastal Current and Florida-Bahamas sampling areas plot separately from the rest of the Atlantic basin (Khim & Krantz, 1996; Surge & Lohmann, 2002). Temporally, the data from a single NASA cruise, SPURS2 (Salinity Processes in the Upper Ocean Regional Study), plots on an independent line from the rest of the Subtropical to Tropical North Pacific. This suggests that the data from this cruise may represent a distinct sampling period with conditions that are not well represented by the majority of the regional data. Furthermore, the Western Pacific region is composed almost entirely of data from two long-term monitoring sites (Manus, Papua New Guinea and Koror, Palau). Although spatial data sets from these sites had similar regression equations relative to temporal data sets, the Palau regression is known to vary interannually, likely due to changes related to the El Niño-Southern Oscillation (Conroy et al., 2017; Morimoto et al., 2002).

4.2. $\delta^{18}O_{sw}$-Salinity Relationships Versus Previous Work

LeGrande and Schmidt (2006) defined a set of 16 $\delta^{18}O_{sw}$-salinity relationships for the mixed layer, constrained by surface ocean circulation, but not explicitly defined in space. Without a clear understanding of region boundaries, it is challenging to draw comparisons between the regions calculated here and the previous data set. However, a selected subset of regions from both, which are likely to have spatial overlap, will be discussed here: Mediterranean Sea, Southern Ocean, and the tropical Pacific region(s). The Mediterranean Sea calculated here is quite similar to that calculated in LeGrande and Schmidt (2006) with slopes at 0.23, 0.28‰/PSU and intercepts at −8.66, −9.24‰, respectively. The similarities between these two regions provide a benchmark for the data throughput from this study, as the machine learning technique here captures the semiclosed region inferred in previous work. The Southern Ocean provides an example of a region with significant increases in data since LeGrande and Schmidt (2006) ($N = 503$ to $N = 1,942$). The addition of data here led to a higher slope that is typical of regions with large freshwater fluxes (0.4 vs. 0.24 ‰/PSU previously) and a lower intercept (−13.75‰ from −8.45‰ previously) reflecting low δ^{18}O runoff and precipitation. This makes a strong case for how additional data in the surface ocean can dramatically change $\delta^{18}O_{sw}$-salinity relationships and shape our interpretations. In the tropics, our data set provides two distinct $\delta^{18}O_{sw}$-salinity relationships: western Pacific and central/eastern Pacific. Splitting of the tropics shows a central/eastern Pacific region with a lower slope, higher intercept compared to the western Pacific region. Our western Pacific region ($y = 0.28x - 9.28$) is quite similar to the fit provided in LeGrande and Schmidt (2006) ($y = 0.27x - 8.88$), which is used to encapsulate the whole of the tropical Pacific in that study. This change in slopes across the tropics is also noted in regional studies of the tropical Pacific (Conroy et al., 2017). The classification of the tropics into two regions provides a clear example of our data set providing additional observed spatial variability compared to previously calculated $\delta^{18}O_{sw}$-salinity regions.

4.3. Applications and Use

4.3.1. Paleoceanography and Paleoclimatology: Coral Forward Modeling

We anticipate that the regions generated in the work will be of specific interest in paleoclimate research that utilizes marine carbonate δ^{18}O records. The linear relationships defined in this work can aid in reconstruction of salinity from proxy δ^{18}O and $\delta^{18}O_{sw}$ records. Additionally, coral δ^{18}O forward modeling, which uses observational $\delta^{18}O_{sw}$ and SST to predict "pseudocoral" time series, requires both regional $\delta^{18}O_{sw}$ data and the locally calibrated slope value to predict the $\delta^{18}O_{sw}$ contribution to coral δ^{18}O records (Reed et al., 2022; Thompson et al., 2011, 2022). In this way, the regional slope indicates whether the coral record should be interpreted as largely responding to $\delta^{18}O_{sw}$ (salinity) or to SST, and the pseudocorals are highly sensitive to changes in this slope (Thompson et al., 2022). The distinction between $\delta^{18}O_{sw}$ or SST dependent between becomes particularly important in locations with large freshwater fluxes such as the western tropical Pacific, where $\delta^{18}O_{sw}$ (salinity) is the main driver of $\delta^{18}O_c$ (Conroy et al., 2017). Using the data set provided here, locally constrained slopes can be used in place of a "bulk tropical Pacific" value when assessing $\delta^{18}O_{sw}$ or SST signals with reduced uncertainty in the slope value, which can cause overestimation of the $\delta^{18}O_{sw}$ contribution (Thompson et al., 2022). Further understanding of the $\delta^{18}O_{sw}$ (salinity) and SST contributions to these proxy records will allow for more detailed assessments of the main controls on the related ENSO signature in these marine carbonate records.

4.3.2. Isotope-Enabled Climate Models

Water movement through the earth system is a critical part of the global climate system and stable water isotopologues provide a unique ability to trace hydroclimate processes in climate models as well as paleoclimate proxies (Dee et al., 2023). Global climate models are increasingly including an isotopic tracer, such as the isotope-enabled Community Earth System Model (iCESM) (Brady et al., 2019). This effort requires a gridded input data set of global $\delta^{18}O_{sw}$ values, such as that computed in this work. Ascertaining the performance of these isotope-enabled climate models is also contingent on accurate modern δ^{18}O measurements. For example, in iCESM (Brady et al., 2019), the simulated $\delta^{18}O_{sw}$-salinity slopes are consistently lower in the Atlantic and Pacific compared to the observation-based slopes provided in LeGrande and Schmidt (2006) and surface ocean δ^{18}O in the Atlantic and Pacific tropics and subtropics is also too depleted. At present, it is unclear whether these differences arise from model biases, creation of regions that are not representative of real-world $\delta^{18}O_{sw}$-salinity relationships, or temporal variability in the $\delta^{18}O_{sw}$-salinity relationship. Regardless, the data set provided here will be useful for future isotope-enabled model benchmarking exercises.

5. Conclusions

Accurate interpretations of marine carbonate δ^{18}O records depend on robust, regional $\delta^{18}O_{sw}$-salinity relationships to understand past changes in ocean salinity, ocean circulation, and climatic variability. Despite limited observational data, we show that neural network learning is able to define regions with specific $\delta^{18}O_{sw}$ and salinity values. With this method, we identify distinct surface ocean regions that range from entire ocean basins to smaller regions associated with unique sources of freshwater forcing, and then calculate the linear $\delta^{18}O_{sw}$-salinity relationships in these regions. We also show that unique secondary slopes are a function of small-scale spatial and temporal variability. The spatial data set developed in this project opens up possibilities to further investigate what drives changes in the $\delta^{18}O_{sw}$-salinity relationship. Additionally, the data processing provided here allows for reproducibility so that regions may be updated as more data becomes available. SOM experiments produced here using atmospheric and ocean tracers show that our E-P and $\delta^{18}O_p$ "atmospheric tracer" produces spatial patterns that are highly similar to the $\delta^{18}O_{sw}$-salinity output used to derived regional fits. This suggests that the $\delta^{18}O_{sw}$-salinity relationship is highly influenced by atmospheric processes globally. Thus, the extra degree of freedom accorded by measuring $\delta^{18}O_{sw}$ along with salinity may provide enhanced ability to find atmospheric signals in modern seawater and paleoseawater archives.

Data Availability Statement

Code and interpolations developed for this research is available at Zenodo (Murray et al., 2023b). Individual data points used in this work, along with their metadata and region assignment, are available at NOAA Paleo Data Search (Murray et al., 2023a).

Acknowledgments
This work was funded by NSF-OCE CAREER 1847791 to JLC. The authors would like to thank all the researchers whose data collection and archiving made this work possible. We would also like to acknowledge Trent Ford, whose work inspired our use of the self-organizing map.

References

Abram, N. J., Wright, N. M., Ellis, B., Dixon, B. C., Wurtzel, J. B., England, M. H., et al. (2020). Coupling of Indo-Pacific climate variability over the last millennium. *Nature*, *579*(7799), 385–392. https://doi.org/10.1038/s41586-020-2084-4

Akhoudas, C. H., Sallée, J.-B., Haumann, F. A., Meredith, M. P., Garabato, A. N., Reverdin, G., et al. (2021). Ventilation of the abyss in the Atlantic sector of the Southern Ocean. *Scientific Reports*, *11*(1), 6760. https://doi.org/10.1038/s41598-021-86043-2

Akman, O., Comar, T., Hrozencik, D., & Gonzales, J. (2019). Data clustering and self-organizing maps in Biology. In *Algebraic and Combinatorial Computational Biology* (pp. 351–374). Elsevier. https://doi.org/10.1016/b978-0-12-814066-6.00011-8

Alibert, C., & McCulloch, M. T. (1997). Strontium/calcium ratios in modern *porites* corals from the Great Barrier Reef as a proxy for sea surface temperature: Calibration of the thermometer and monitoring of ENSO. *Paleoceanography*, *12*(3), 345–363. https://doi.org/10.1029/97PA00318

Anderson, R. F., Mawji, E., Cutter, G. A., Measures, C. I., & Jeandel, C. (2014). Geotraces: Changing the way we explore ocean chemistry. *Oceanography*, *27*(1), 50–61. https://doi.org/10.5670/oceanog.2014.07

Bauch, D., Erlenkeuser, H., & Andersen, N. (2005). Water mass processes on Arctic shelves as revealed from $\delta^{18}O$ of H_2O. *Global and Planetary Change*, *48*(1–3), 165–174. https://doi.org/10.1016/j.gloplacha.2004.12.011

Bauch, D., Hölemann, J., Willmes, S., Gröger, M., Novikhin, A., Nikulina, A., et al. (2010). Changes in distribution of brine waters on the Laptev Sea shelf in 2007. *Journal of Geophysical Research*, *115*, C11008. https://doi.org/10.1029/2010JC006249

Benway, H. M., & Mix, A. C. (2004). Oxygen isotopes, upper-ocean salinity, and precipitation sources in the eastern tropical Pacific. *Earth and Planetary Science Letters*, *224*(3–4), 493–507. https://doi.org/10.1016/j.epsl.2004.05.014

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Chen, T., Cobb, K. M., Roff, G., Zhao, J., Yang, H., Hu, M., & Zhao, K. (2018). Coral-derived western Pacific tropical sea surface temperatures during the last millennium. *Geophysical Research Letters*, *45*, 3542–3549. https://doi.org/10.1002/2018GL077619

Conroy, J. L., Murray, N. K., Patterson, G. S., Schore, A. I. G., Ikuru, I., Cole, J. E., et al. (2023). Equatorial Undercurrent influence on surface seawater $\delta^{18}O$ values in the Galápagos. *Geophysical Research Letters*, *50*, e2022GL102074. https://doi.org/10.1029/2022GL102074

Conroy, J. L., Thompson, D. M., Cobb, K. M., Noone, D., Rea, S., & Legrande, A. N. (2017). Spatiotemporal variability in the $\delta^{18}O$-salinity relationship of seawater across the tropical Pacific ocean: Isotope-salinity relationships. *Paleoceanography*, *32*, 484–497. https://doi.org/10.1002/2016PA003073

Craig, H., & Gordon, L. I. (1965). Deuterium and oxygen 18 variations in the ocean and the marine atmosphere. In *Stable Isotopes in Oceanographic Studies and Paleotemperatures* (pp. 9–130).

Dee, S. G., Bailey, A., Conroy, J. L., Atwood, A., Stevenson, S., Nusbaumer, J., & Noone, D. (2023). Water isotopes, climate variability, and the hydrological cycle: Recent advances and new frontiers. *Environmental Research: Climate*, *2*(2), 022002. https://doi.org/10.1088/2752-5295/accbe1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Delaygue, G., Bard, E., Rollion, C., Jouzel, J., Stiévenard, M., Duplessy, J.-C., & Ganssen, G. (2001). Oxygen isotope/salinity relationship in the northern Indian Ocean. *Journal of Geophysical Research*, *106*(C3), 4565–4574. https://doi.org/10.1029/1999JC000061

Dubinina, E. O., Kossova, S. A., Miroshnikov, A. Y., & Kokryatskaya, N. M. (2017). Isotope (δD, $\delta^{18}O$) systematics in waters of the Russian Arctic seas. *Geochemistry International*, *55*(11), 1022–1032. https://doi.org/10.1134/S0016702917110052

Dubinina, E. O., Miroshnikov, A. Y., Kossova, S. A., & Shchuka, S. A. (2019). Modification of the Laptev Sea freshened shelf waters based on isotope and salinity relations. *Geochemistry International*, *57*, 1–19. https://doi.org/10.1134/S001670291901004X

Durack, P. J., Wijffels, S. E., & Matear, R. J. (2012). Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*, *336*(6080), 455–458. https://doi.org/10.1126/science.1212222

Epstein, S., Buchsbaum, R., Lowenstam, H. A., & Urey, H. C. (1953). Revise carbonate-water isotopic temperature scale. *Geological Society of America Bulletin*, *64*(11), 1315. https://doi.org/10.1130/0016-7606(1953)64[1315:RCITS]2.0.CO;2

Gagan, M. K., Ayliffe, L. K., Hopley, D., Cali, J. A., Mortimer, G. E., Chappell, J., et al. (1998). Temperature and surface-ocean water balance of the Mid-Holocene tropical western Pacific. *Science*, *279*(5353), 1014–1018. https://doi.org/10.1126/science.279.5353.1014

Gorman, M. K., Quinn, T. M., Taylor, F. W., Partin, J. W., Cabioch, G., Austin, J. A., Jr., et al. (2012). A coral-based reconstruction of sea surface salinity at Sabine Bank, Vanuatu from 1842 to 2007 CE. *Paleoceanography*, *27*, PA3226. https://doi.org/10.1029/2012PA002302

Grothe, P. R., Cobb, K. M., Liguori, G., Di Lorenzo, E., Capotondi, A., Lu, Y., et al. (2020). Enhanced El Niño-Southern Oscillation variability in recent decades. *Geophysical Research Letters*, *47*, e2019GL083906. https://doi.org/10.1029/2019GL083906

Holbourn, A., Kuhnt, W., Clemens, S. C., & Heslop, D. (2021). A ~12 Myr Miocene record of East Asian Monsoon variability from the South China sea. *Paleoceanography and Paleoclimatology*, *36*, e2021PA004267. https://doi.org/10.1029/2021PA004267

Khim, B. K., & Krantz, D. E. (1996). Oxygen isotopic identity of the Delaware coastal current. *Journal of Geophysical Research*, *101*(C7), 16509–16514. https://doi.org/10.1029/96JC00201

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69. https://doi.org/10.1007/BF00337288

Konecky, B. L., Noone, D. C., & Cobb, K. M. (2019). The influence of competing hydroclimate processes on stable isotope ratios in tropical rainfall. *Geophysical Research Letters*, *46*, 1622–1633. https://doi.org/10.1029/2018GL080188

Krief, S., Hendy, E. J., Fine, M., Yam, R., Meibom, A., Foster, G. L., & Shemesh, A. (2010). Physiological and isotopic responses of scleractinian corals to ocean acidification. *Geochimica et Cosmochimica Acta*, *74*(17), 4988–5001. https://doi.org/10.1016/j.gca.2010.05.023

LeGrande, A. N., & Schmidt, G. A. (2006). Global gridded data set of the oxygen isotopic composition in seawater. *Geophysical Research Letters*, *33*, L12604. https://doi.org/10.1029/2006GL026011

Liu, Y., & Weisberg, R. H. (2011). A review of self-organizing map applications in Meteorology and Oceanography. In J. I. Mwasiagi (Ed.), *Self organizing maps—Applications and novel algorithm design*. InTech. https://doi.org/10.5772/13146

Liu, Y., Weisberg, R. H., & Mooers, C. N. K. (2006). Performance evaluation of the self-organizing map for feature extraction. *Journal of Geophysical Research*, *111*, C05018. https://doi.org/10.1029/2005JC003117

Mitsuguchi, T., Matsumoto, E., Abe, O., Uchida, T., & Isdale, P. J. (1996). Mg/Ca thermometry in coral skeletons. *Science*, *274*(5289), 961–963. https://doi.org/10.1126/science.274.5289.961

Moosavi, V., Packmann, S., & Vallés, I. (2014). *SomPy: A Python library for self organizing map (SOM)*. Github. Retrieved from https://github.com/sevamoo/SOMPY

Morimoto, M., Abe, O., Kayanne, H., Kurita, N., Matsumoto, E., & Yoshida, N. (2002). Salinity records for the 1997-98 El Nino from western Pacific corals. *Geophysical Research Letters*, *29*(11), 1540. https://doi.org/10.1029/2001GL013521

Murray, N. K., Munoz, A. R., & Conroy, J. L. (2023a). NOAA/WDS Paleoclimatology—Global surface ocean δ18O-salinity relationship data for paleoclimatic reconstruction [Dataset]. National Centers for Environmental Information. https://doi.org/10.25921/zvtv-qj95.NOAA

Murray, N. K., Munoz, A. R., & Conroy, J. L. (2023b). Surface ocean δ18O-salinity regions as revealed by machine learning [Software]. Zenodo. https://doi.org/10.5281/zenodo.6557204

Nan, Y., Tian, F., Hu, H., Wang, L., & Zhao, S. (2019). Stable isotope composition of river waters across the world. *Water*, *11*(9), 1760. https://doi.org/10.3390/w11091760

O'Neil, J. R., Clayton, R. N., & Mayeda, T. K. (1969). Oxygen isotope fractionation in divalent metal carbonates. *The Journal of Chemical Physics*, *51*(12), 5547–5558. https://doi.org/10.1063/1.1671982

Reed, E. V., Thompson, D. M., & Anchukaitis, K. J. (2022). Coral-based sea surface salinity reconstructions and the role of observational uncertainties in inferred variability and trends. *Paleoceanography and Paleoclimatology*, *37*, e2021PA004371. https://doi.org/10.1029/2021PA004371

Reverdin, G., Waelbroeck, C., Pierre, C., Akhoudas, C., Aloisi, G., Benetti, M., et al. (2022). The CISE-LOCEAN sea water isotopic database (1998–2021). *Earth System Science Data Discussions*, *14*(6), 2721–2735. https://doi.org/10.5194/essd-14-2721-2022

Robbins, L. L., Knorr, P. O., Wynn, J. G., Hallock, P., & Harries, P. J. (2017). Interpreting the role of pH on stable isotopes in large benthic foraminifera. *ICES Journal of Marine Science*, *74*(4), 955–964. https://doi.org/10.1093/icesjms/fsw056

Rohling, E. J. (2013). Paleoceanography, physical and chemical proxies: Oxygen isotope composition of seawater. In *Encyclopedia of quaternary science* (pp. 915–922). Elsevier. https://doi.org/10.1016/B978-0-444-53643-3.00293-4

Russon, T., Tudhope, A. W., Hegerl, G. C., Collins, M., & Tindall, J. (2013). Inter-annual tropical Pacific climate variability in an isotope-enabled CGCM: Implications for interpreting coral stable oxygen isotope records of ENSO. *Climate of the Past*, *9*(4), 1543–1557. https://doi.org/10.5194/cp-9-1543-2013

Schlegel, R. W., Oliver, E. C. J., Perkins-Kirkpatrick, S., Kruger, A., & Smit, A. J. (2017). Predominant atmospheric and oceanic patterns during coastal marine Heatwaves. *Frontiers in Marine Science*, *4*, 323. https://doi.org/10.3389/fmars.2017.00323

Schmidt, G. A. (1999). Global seawater oxygen-18 database. Retrieved from https://data.giss.nasa.gov/o18data/

Schmidt, G. A., Bigg, G. R., & Rohling, E. J. (1999). Lobal seawater oxygen-18 database—v1.22. Retrieved from https://data.giss.nasa.gov/o18data/

Spero, H. J., Bijma, J., Lea, D. W., & Bemis, B. E. (1997). Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes. *Nature*, *390*(6659), 497–500. https://doi.org/10.1038/37333

Stevenson, S., Powell, B., Cobb, K. M., Nusbaumer, J., Merrifield, M., & Noone, D. (2018). Twentieth century seawater δ^{18}O dynamics and implications for coral-based climate reconstruction. *Paleoceanography and Paleoclimatology*, *33*(6), 606–625. https://doi.org/10.1029/2017PA003304

Stevenson, S., Powell, B. S., Merrifield, M. A., Cobb, K. M., Nusbaumer, J., & Noone, D. (2015). Characterizing seawater oxygen isotopic variability in a regional ocean modeling framework: Implications for coral proxy records. *Paleoceanography*, *30*, 1573–1593. https://doi.org/10.1002/2015PA002824

Surge, D. M., & Lohmann, K. C. (2002). Temporal and spatial differences in salinity and water chemistry in SW Florida estuaries: Effects of human-impacted watersheds. *Estuaries*, *25*(3), 393–408. https://doi.org/10.1007/BF02695982

Thompson, D. M., Ault, T. R., Evans, M. N., Cole, J. E., & Emile-Geay, J. (2011). Comparison of observed and simulated tropical climate trends using a forward model of coral $\delta\ ^{18}$O. *Geophysical Research Letters*, *38*, L14706. https://doi.org/10.1029/2011GL048224

Thompson, D. M., Conroy, J. L., Konecky, B. L., Stevenson, S., DeLong, K. L., McKay, N., et al. (2022). Identifying hydro-sensitive coral δ 18O records for improved high-resolution temperature and salinity reconstructions. *Geophysical Research Letters*, *49*, e2021GL096153. https://doi.org/10.1029/2021GL096153

Yu, L., Josey, S. A., Bingham, F. M., & Lee, T. (2020). Intensification of the global water cycle and evidence from ocean salinity: A synthesis review. *Annals of the New York Academy of Sciences*, *1472*(1), 76–94. https://doi.org/10.1111/nyas.14354

References From the Supporting Information

Melnichenko, O., Hacker, P., Maximenko, N., Lagerloef, G., & Potemra, J. (2016). Optimal interpolation of Aquarius sea surface salinity. *Journal of Geophysical Research: Oceans*, *121*, 602–616. https://doi.org/10.1002/2015JC011343