esa

ECOSPHERE

EMERGING TECHNOLOGIES

The hunt for red tides: Deep learning algorithm forecasts shellfish toxicity at site scales in coastal Maine

Isabella Grasso,[1] Stephen D. Archer,[1] Craig Burnell,[1] Benjamin Tupper,[1] Carlton Rauschenberg,[1] Kohl Kanwit,[2] and Nicholas R. Record[1,†]

[1]*Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine, USA*
[2]*Maine Department of Marine Resources, East Boothbay, Maine, USA*

Citation: Grasso, I., S. D. Archer, C. Burnell, B. Tupper, C. Rauschenberg, K. Kanwit, and N. R. Record. 2019. The hunt for red tides: Deep learning algorithm forecasts shellfish toxicity at site scales in coastal Maine. Ecosphere 10(12):e02960. 10.1002/ecs2.2960

Abstract. Farmed and wild harvest shellfish industries are increasingly important components of coastal economies globally. Disruptions caused by harmful algal blooms (HABs), colloquially known as red tides, are likely to worsen with increasing aquaculture production, environmental pressures of coastal development, and climate change, necessitating improved HAB forecasts at finer spatial and temporal resolution. We leveraged a dataset of chemical analytical toxin measurements in coastal Maine to demonstrate a new machine learning approach for high-resolution forecasting of paralytic shellfish toxin accumulation. The forecast used a deep learning neural network to provide weekly site-specific forecasts of toxicity levels. The algorithm was trained on images constructed from a chemical fingerprint at each site composed of a series of toxic compound measurements. Under various forecasting configurations, the forecast had high accuracy, generally >95%, and successfully predicted the onset and end of nearly all closure-level toxic events at the site scale at a one-week forecast time. Tests of forecast range indicated a decline in accuracy at a three-week forecast time. Results indicate that combining chemical analytical measurements with new machine learning tools is a promising way to provide reliable forecasts at the spatial and temporal scales useful for management and industry.

Key words: forecast; harmful algal bloom; Maine; neural network; paralytic shellfish toxin.

Received 14 March 2019; revised 9 August 2019; accepted 8 October 2019. Corresponding Editor: Bistra Dilkina.

† **E-mail:** nrecord@bigelow.org

Introduction

Farmed and wild harvest shellfish industries are important components of coastal economies globally. Although aquaculture is one of the fastest growing food production sectors in the world, every year shellfish aquaculture and fisheries are disrupted by losses associated with the bioaccumulation of toxins generated by harmful algal blooms (HABs), known colloquially as red tides. The annual economic impact of HABs in the United States alone averages $50 million a year, including direct losses to the shellfish industries, the costs of protecting public health through monitoring programs and fisheries management, and the economic costs to tourism and recreation (Anderson et al. 2000). The disruption caused by HABs is likely to worsen with increasing aquaculture production, environmental pressures of coastal development, and climate change (De Silva and Soto 2009). To meet these growing challenges, we will need improved methods for forecasting HABs at finer spatial and temporal scales than are currently available.

The Department of Marine Resources in Maine, USA, has recently launched a new monitoring process that incorporates a chemical analytical approach, producing a substantial and largely untapped dataset for testing predictive methods. In the present study, we leveraged this dataset to demonstrate a new machine learning approach for high-resolution forecasting of paralytic shellfish toxin (PST) accumulation in coastal Maine.

The primary causative organism of HABs in the Gulf of Maine, *Alexandrium catenella* (*Alexandrium fundyense*), is a dinoflagellate that is actively motile in the water column in the spring and summer as an asexual vegetative cell, transforms to sexual gametes in late summer and autumn, that then fuse to form a resting cyst phase that persists in the sediment through the winter (Yentsch et al. 1980). *Alexandrium catenella* is responsible for the synthesis of the potent neurotoxin saxitoxin (STX) and its derivatives, collectively comprising the PSTs, that are bioaccumulated by shellfish and present a potentially lethal risk to humans and other higher trophic level consumers (Stüken et al. 2011, Cusick and Sayler 2013). To protect the public, monitoring of toxicity in shellfish along the Maine coastline has been conducted by the Department of Marine Resources since 1958, with a more comprehensive program introduced in 1975 following gulf-wide PSP events in 1972 and 1974 (Shumway et al. 1988, Bean et al. 2005). *Mytilus edulis* is used as an indicator species because of its abundance and ability to filter feed in generally colder waters than other shellfish, which leads these mussels to become toxic before other species (Bean et al. 2005). Closures of areas of the coastline to shellfish harvesting are initiated when the PST levels in the tissue of *M. edulis* approach or reach a level of 80 μg of saxitoxin equivalence 100 g^{-1} of tissue.

Alexandrium catenella is a challenging HAB species for forecasters. There is a wide variety of approaches to forecasting the development and toxicity of HABs (reviewed in Anderson et al. 2015, Davidson et al. 2016). Coupled biological and physical models have been developed to forecast the development of phytoplankton blooms that include toxic species and to usefully predict the advection of HAB-containing waters into shellfish harvesting regions (e.g., Escalera et al. 2010, McGillicuddy et al. 2011). Such ecosystem forecasts are potentially a valuable tool for management of closures and for industry response, helping to guide monitoring, making decisions proactive, and helping to buffer economic swings. However, there are certain challenges to making forecasts useful. For example, there is not a consistent relationship between interannual variations in *A. catenella* abundance and shellfish toxicity (McGillicuddy et al. 2005, Apurva et al. 2013). Moreover, because of uncertainties around the mechanisms used in models, forecasts are much more useful at seasonal or interannual timescales and at broad spatial scales than at weekly or site-specific scales (Stumpf et al. 2009, Anderson et al. 2014). For toxic species that occur in sufficient abundance, satellite imagery can be used to assess bloom extent and to track bloom movement in near-real time (Amin et al. 2009, Shutler et al. 2012). In the case of the extensive blooms containing the toxic *Karenia brevis* in the Gulf of Mexico, for example, satellite imagery has been used in combination with field and meteorological data and numerical models to provide near-real-time information on the location of toxic blooms and to assess their potential to be advected into near-shore regions (Hu et al. 2016). In contrast, *A. catenella* occurs at relatively low abundance (<1000 cell/L) and comprises a small proportion of the total phytoplankton signal, precluding satellite-based predictive approaches. From the industry and direct management point of view, broad spatial and temporal scales are less useful for decision making. Growers, harvesters, and managers are making staffing, effort, harvesting, monitoring, business, and safety decisions at much shorter timescales and need information at site-specific levels and at weekly timescales.

Neural networks have shown some potential for forecasting at weekly timescales (Velo-Suárez and Gutiérrez-Estrada 2007), and deep learning algorithms have revolutionized machine learning in recent years (Liu et al. 2017, Chollet and Allaire 2018). Machine learning relies on a series of matched inputs and outputs to train a program to develop empirical relationships that can be utilized to predict future outputs. This process has been improved upon through deep learning by creating a deeper structure of connections that allows for the completion of more complex tasks such as image recognition and language

translation (Chollet and Allaire 2018). This key improvement allows algorithms to better grapple with the complexity of an ecosystem. We used four years of toxin data, measured approximately weekly at sites across coastal Maine and configured to mimic an image classification task, to test the ability of deep learning to forecast toxicity levels at finer spatial and temporal scales.

Methods

We utilized data generated weekly by Bigelow Analytical Services (BAS) and the Department of Marine Resources (DMR) at harvesting sites on the Maine coastline (Fig. 1; ~3000 samples). Bigelow Analytical Services toxin testing follows a chemical analytical technique that uses high-performance liquid chromatography (HPLC; Rourke et al. 2008, Van De Riet et al. 2009) for monitoring PST levels (Anon 2011). Through each season, *M. edulis* and other shellfish species were collected weekly by DMR from each location and delivered to BAS for analysis. Tissue samples were homogenized and extracted from a minimum of 12 animals, and 5 g was used for the extraction process. For each sample, two HPLC analyses were utilized, one for the carbamate, decarbamoyl, and N-sulfocarbamoyl PST congeners, and the second to analyze the N-sulfocarbamoyl congeners. Agilent 1260 Infinity HPLCs with fluorescence detection were connected to Analytical Scientific Instruments US, Model 310 Post-Column Reactors. At the height of the PST season, a total of four HPLC systems were run almost continuously in order to process roughly 50 samples each day. Each sample was compared to the NRC-CNRC obtained PST standards in order to identify each STX congener present based on retention time. To calculate a total PST toxicity for a sample, toxicity equivalence factors relative to STX were assigned to each compound based on tests that determined the acute response to intraperitoneal injections in mice. The total STX-eq. toxin scores for each sample were reported to DMR within 24 h and were used to make decisions on the risks to public safety and closure or opening of areas for shellfish harvesting.

We split the toxicity scale into four classes, where the highest was the National Shellfish Sanitation Program's established cutoff limit for closure, 80 μg equivalence 100 g^{-1} of shellfish tissue. The classification categories 0–3, respectively, were 0–10 μg 100 g^{-1} shellfish (80% of samples), 10–30 μg 100 g^{-1} (11% of samples), 30–80 μg 100 g^{-1} (5% of samples), and >80 μg 100 g^{-1} (4% of samples). This classification approach followed a common approach of categorizing by severity for forecasting (e.g., hurricanes, Simpson 1971; HAB spatial extent Kleindinst et al. 2014), which facilitates proactive management. While the objective is the binary distinction between closure and non-closure-level events, addition of two intermediate classification levels helped to address the data imbalance (>90% non-closures).

We used the Keras open-source package in R, which is a wrapper for an open-source Python library (Allaire and Chollet 2018). Keras is a neural network application programming interface, which is being run on top of TensorFlow in this study. It allows users to operate convolutional neural networks. In particular, we utilized the Keras sequential model, which allows users to build a neural network layer by layer (Chollet et al. 2015). The model includes an input layer with dropout, a fully connected hidden layer with dropout, and an output layer. Dropout is a technique used to avoid overfitting, where a fixed probability of dropping each unit is determined and the choice of units is random (Srivastava et al. 2014). In this context, dropping a unit means deleting that unit as well as all of its connections during training (Srivastava et al. 2014). To determine the dropout rate, we used the R package *tfruns*, which is designed for hyperparameter optimization (Abadi et al. 2015). It allows users to track the performance of several training runs by defining flags for parameters and training over all possible combinations of flag values (Abadi et al. 2015). Flags were created for dropout, batch size, units, and epochs, and final values were decided by validation and testing accuracy levels (i.e., the fraction of correct classifications). Other decisions included optimizer and loss function. The optimizer used was Adam, which is a stochastic optimization method favored for problems with a lot of noise, which is true for most problems in computational ecology (Kingma and Ba 2014). The loss function used was categorical cross-entropy, which is generally used when the output is defined by several possible categories.

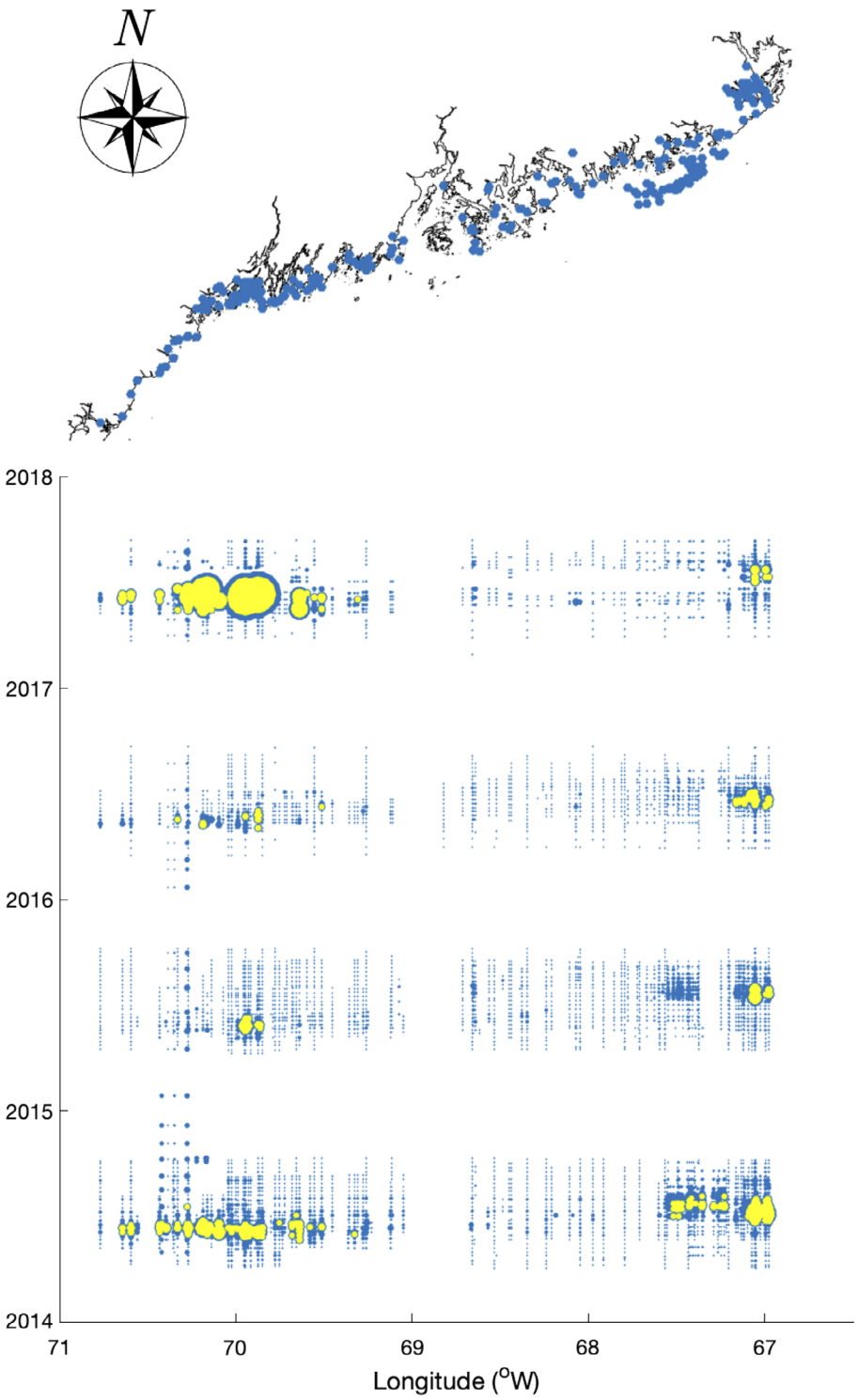

Fig. 1. Monitoring sites and seasonal shellfish toxicity dynamics along coastal Maine. The size of the circle scales with total paralytic shellfish toxin toxicity. The yellow circles indicate closure-level toxicity measurements.

The data were restructured as a list of images against their classification categories. Each sample at each site was associated with a constructed image, or two-dimensional array, representing past conditions. Each sample contains the toxicity level of the twelve toxins we tested for. Each image then contains this information for five samples, so an image contains twelve rows, one for each toxin, and five columns, one for each sample (i.e., going back five weeks). The images were filtered to remove any with gaps longer than ten days between samples, any without five previous samples, and any with missing data, and the remaining set of images was normalized to the maximum value so that the full range of measurements spanned a 0–1 scale. This resulted in 3198 good images, with 2551 labeled 0, 354 labeled 1, 161 labeled 2, and 132 labeled 3. Each image contained the toxicity levels of the twelve toxins for the past five weeks, that is, the chemical fingerprint associated with the current sample (Fig. 2). Thus, data from previous weeks (i.e., the image) could train the algorithm to forecast the subsequent week's (or weeks') toxicity classification category.

A validation set of 20% of training data was used to tune the model parameters, while the testing set was used as an unbiased measure of the model performance. The metrics of both the validation and testing sets were tracked. In forecasting mode, the forecast year served as validation and was not used in the training. We trained and tested the data in multiple configurations. First, we used three of the four years for training and the fourth year for testing for each of the four years of data. We also ran in a simulated forecasting mode, where only previous years were available for training. That is, for 2015 forecasts, only 2014 data were used in the training; for 2016, 2014–2015 data were used in the training; for 2017 forecasts, 2014–2016 data were used in the training. Under these forecasting configurations, forecasting at a particular time and site required five previous weeks of data to generate the image and would provide a forecast one week in advance. Finally, to test forecast range, we extended this lag from one to ten weeks, using the same approach to algorithm training and testing, and evaluated how forecast skill declined with increasing lag time. In each configuration, ensembles of 50 forecasts were run—with each ensemble member being a randomized initialization of the neural network—and the mean and variance were used to characterize accuracy for each configuration.

Results

The *tfruns* optimization resulted in a total model size of 12,740 parameters, with 3904 parameters in the input layer, 8320 in the hidden layer, and 516 in the output layer. Hyperparameter values were either optimized or chosen based on Keras specifications (Table 1). The validation split proportion follows the commonly used 80/20 rule, following the Pareto principle (Chollet and Allaire 2018). The number of output units must match the output shape, which in our case is four because we have four possible labels (Chollet et al. 2015). Under this configuration, forecast loss scores converged quickly (Fig. 3).

Toxicity levels have a general seasonal pattern, with high levels in the spring and summer months (Fig. 1). Closure-level toxicity events occurred earlier in the season in western Maine and later in the season in Down East Maine in all four years, consistent with the generally earlier warming in the Western Maine Coastal Current (WMCC; Thomas et al. 2017), though there was considerable variability in the phenology, duration, extent, and magnitude of these events. For example, in 2016, which was the second warmest year recorded for sea surface temperature in the Gulf of Maine, there were comparatively few closure-level toxicity events. For the four years, the number of closure-level measurements was 149, 20, 34, and 120, respectively. There was also notable spatial variability. In western Maine, closure-level events occurred from the western border with New Hampshire across roughly the western third of the coast. This stretch of the coast is influenced by the WMCC, which flows from east to west and varies seasonally and interannually depending largely on outflow from the Penobscot river, but is generally more influenced by the warmer western gyre within the Gulf of Maine. Midcoast Maine, where the Penobscot River empties, had no closure-level measurements, consistent with the region being termed the PSP Sandwich because of the general lack of toxicity, in contrast to eastern and western Maine on either side (Shumway et al. 1988). In eastern

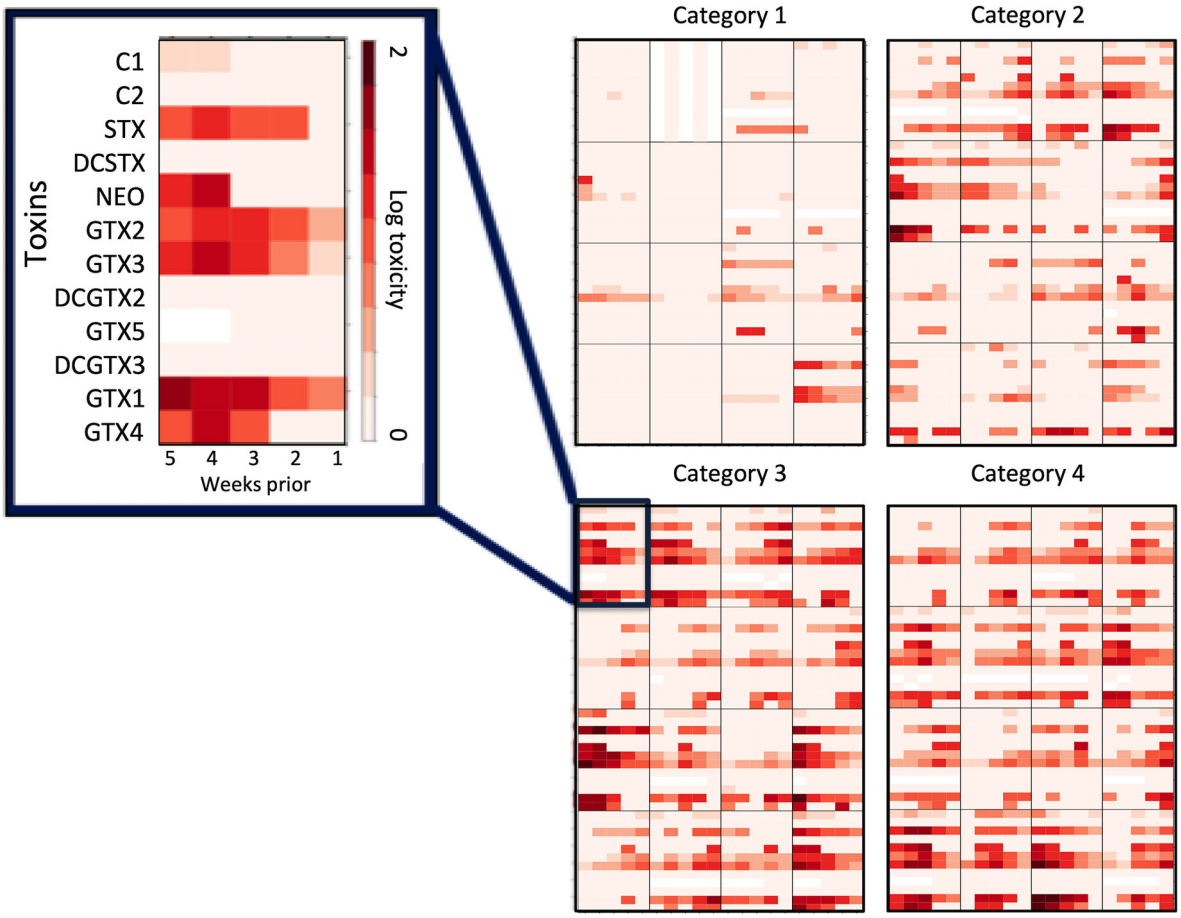

Fig. 2. Examples of toxicity images used for algorithm training. Each image, comprising the fingerprint, is composed of 12 toxins (rows) measured over 5 weeks (columns) prior to the target forecast date. The images are then classified into the four toxicity levels based on the total toxicity measurement in the subsequent week (or at a greater lag, see Methods).

Table 1. Optimized/chosen parameter values as determined by *tfruns* or Keras documentation.

Parameters	Value options	Optimized/chosen values
Dropout rate (input layer)	0.1, 0.2, 0.3, 0.4	0.3
Dropout rate (hidden layer)	0.1, 0.2, 0.3, 0.4	0.1
Batch size	8, 16, 32, 64	8
Number of units (input layer)	64, 128, 256	64
Number of units (hidden layer)	32, 64, 128	128
Number of units (output layer)	N/A, chosen based on documentation	4
Number of epochs	5, 10, 15, 20	20
Validation split proportion	N/A, chosen based on documentation	0.2

Maine, closure-level measurements occurred in roughly the easternmost quarter of the coast. This stretch of the coast is influenced by the Eastern Maine Coastal Current, which flows from east to west, is supplied by the colder Labrador Current, and often exhibits dynamics that are disconnected with the western portion of the coast.

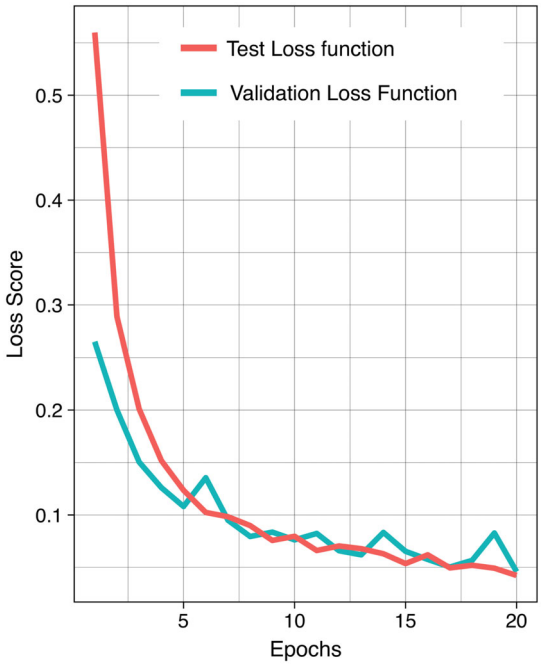

Fig. 3. Example of loss score convergence, for 2017 forecasting run. Loss score as a function of number of epochs of both testing set and validation set of optimized model.

There is some degree of temporal autocorrelation in the dataset, which can be seen by comparing the mean prior toxicity level to subsequent toxicity categories (Fig. 4A). This relationship, however, is far too weak to form the basis for a reliable forecast. Using a threshold toxicity value would either have a very high false-positive rate, if set low, or a very high false-negative rate, if set high. In contrast, the neural network forecast performed extremely well at a one-week advance notice time frame (Fig. 4B–D). Forecast accuracy was very high both when tuned using outside years and in the simulated forecasting mode (Table 2). Importantly, the algorithm correctly forecasted nearly every closure-level event. In 2017, for example, using 2014–2016 as training data, all closure-level events were correctly forecasted with no false positives in this category. Conversely, the algorithm also predicted when the closure-level event would end in all cases, with no false negatives in 2017. In prior years, the forecast performed nearly as well. The year 2014 had the lowest forecast accuracy, though still quite high, and still with very few closure-level events missed. This year also had the highest number of closure-level events (142). To evaluate a measure of forecast confidence, we computed the logit score for each individual forecast. For accurate forecasts (true positive and true negative of closure-level toxicity), logit scores were much higher than for inaccurate forecasts (Fig. 5). This suggests that a low-confidence value for a particular forecast could

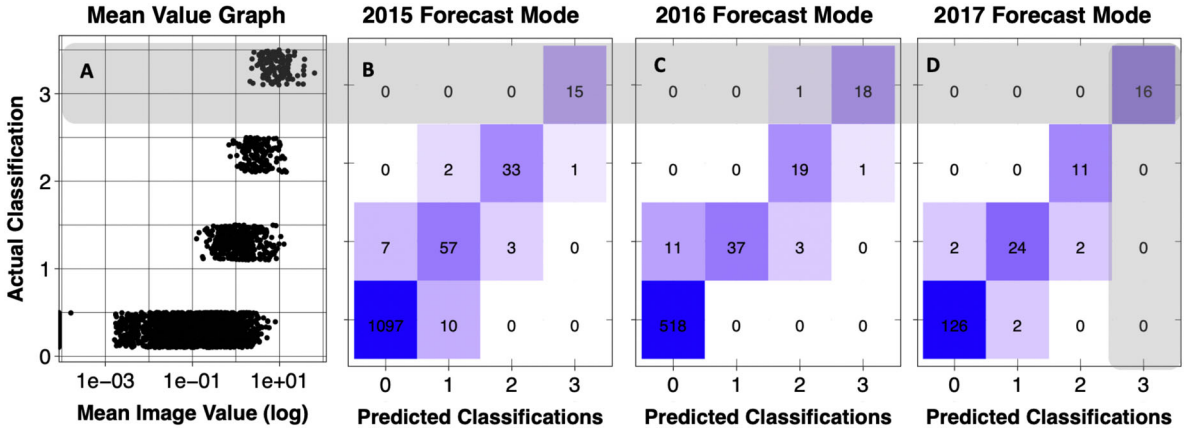

Fig. 4. Forecast algorithm results. (A) A naive model, for comparison, that uses the mean toxicity value to predict the subsequent week's toxicity level. (B–D) Forecasts from the neural network forecast run in simulated true forecasting mode, that is, using prior years as training data and using as test data the measurements from 2015 (B), 2016 (C), and 2017 (D). Random scatter around classification categories is added for visualization in (A). Classification level 3 indicates closure-level toxicity.

Table 2. Model performance by configuration. In the simulated forecasting mode, only prior years were used in training.

Test year	Accuracy (%) training using other three years	Accuracy (%) simulated forecasting mode
2014	95.3 ± 0.2	
2015	99.0 ± 0.1	98.0 ± 0.1
2016	97.2 ± 0.1	97.1 ± 0.1
2017	96.1 ± 0.2	96.1 ± 0.2

be used as an indicator of potential overestimate or underestimate of toxicity.

To test the forecast range, we used increasing lags between the measured data and the target forecast conditions, testing the ability to forecast from one to ten weeks in advance. Forecast accuracy was essentially unchanged moving from a range of one week to a range of two weeks, but at three weeks and longer, accuracy dropped sharply for all tests (Fig. 6). Following the drop, accuracy leveled off for each year's test. There was substantial year-to-year variability in accuracy following this drop, with accuracy ranging from ~70% to 90%. These numbers represent essentially the accuracy that could be achieved in each year by simply always forecasting the lowest classification category, which comprised roughly 70–90% of the data, depending on the model run.

Discussion

All ecosystem forecasting programs must contend with monitoring costs, computational limitations, and inherent stochasticity. The traditional approach of using biological–physical coupled models currently has more utility at regional spatial scales and seasonal or interannual timescales, for many of these reasons.

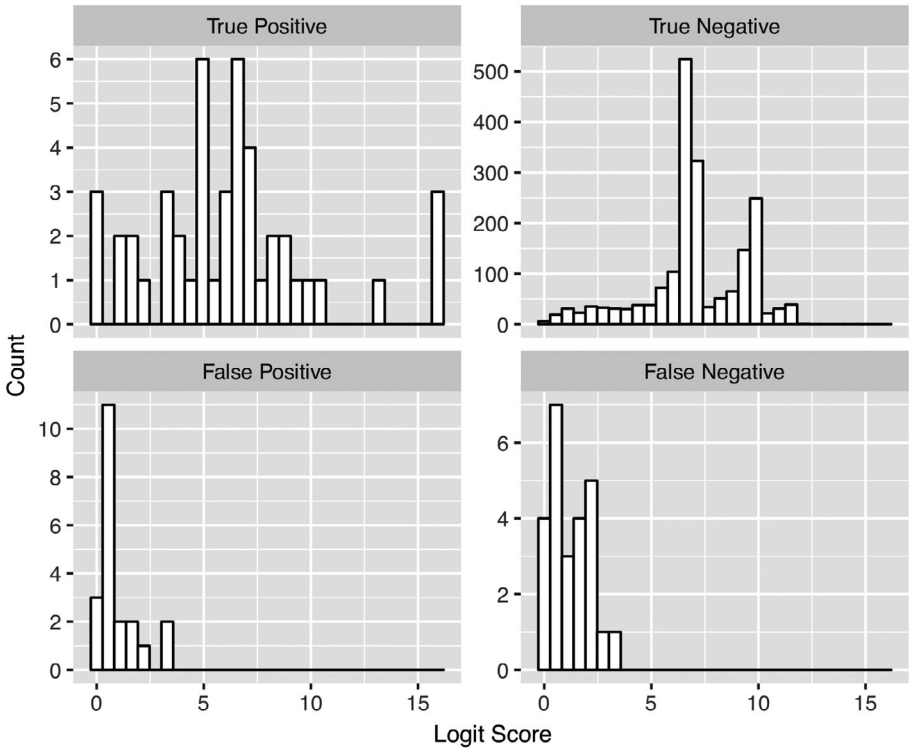

Fig. 5. Histograms of logit scores for all forecasts made in simulated forecasting mode. True positives represent correctly forecasted closure-level toxicity; true negatives represent correctly forecasted sub-closure-level toxicity; false positives represent forecasts that overestimate toxicity; false negatives represent forecasts that underestimate toxicity.

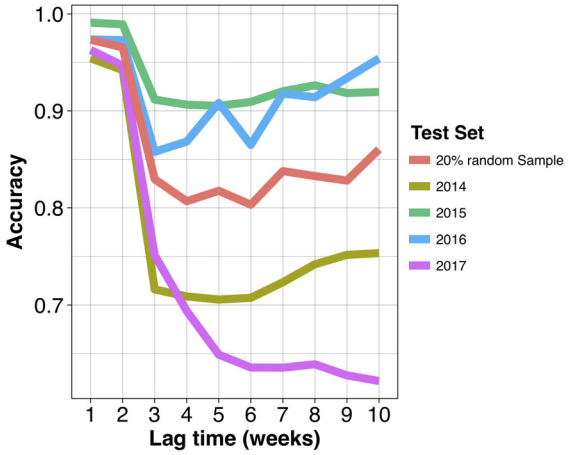

Fig. 6. Accuracy as a function of forecast range. Neural network forecast of each year using the other years as training data, with forecast extended from one to ten weeks in advance. Each year's test used the other years as training.

Machine learning can offer creative ways forward where biological–physical modeling finds limitations (Record et al. 2018). For forecasts that operate at finer spatial (site-specific) and temporal (~weekly) scales, there is a need to explore new methods. This scale has particular utility for growers and harvesters, who use information specific to their sites to make decisions often at week-to-week timescales. Managers concerned with public safety are focused similarly on these time and spatial scales. Conventional forecasts have had limited utility at this scale.

The design of this forecasting system helps to overcome some of the barriers in providing fine-scale forecasts. The application of deep learning helps to incorporate the high-variance information contained in local, weekly measurements, allowing the forecasts to operate at fine scales. Neural networks have been used already to some degree of success in this regard (Velo-Suárez and Gutiérrez-Estrada 2007, Kang et al. 2011, Guallar et al. 2016, Qin et al. 2017, Wang et al. 2019). However, these systems have generally focused on forecasting cell counts or biomass of target species, or total chlorophyll concentrations. There is often a decoupling between abundance measures and toxicity of target species. By incorporating toxic compound measurements into the forecasting algorithm directly, we provided a much more information-rich dataset for the neural network, and we trained it on the state variable (i.e., toxicity) that is of direct relevance to industry and public safety management.

The major drawback to neural networks, as with many basically empirical approaches, is opacity. It can be difficult to access or interpret the underlying correlations, which can be complex and subtle, making the system potentially sensitive to nonstationarity biases. If the underlying relationships driving the predictions change drastically, the forecast can fail. Thus, one must apply caution when extending the forecast to a larger region, a new environment, or a shifting environmental regime. Merging conventional deterministic models with empirically driven machine learning approaches is one way to draw mechanistic links between oceanographic processes and the toxicity measurements collected by managers. Overcoming opacity is an emerging priority in machine learning research (Burrell 2016) and should be considered as these tools are increasingly used in ecological applications.

Conclusion

Neural networks offer advantages over conventional ecosystem forecasting techniques, particular at fine spatial and temporal scales. For HABs in coastal Maine, the forecasting algorithm was able to reliably predict closure-level toxic events at a two-week advance notice and at the scale of individual sites. This forecasting scale is well attuned to the needs of industry and management. By operating directly on toxicity data, the forecast incorporates information directly relevant to decision makers in industry and management. There is potential to merge this machine learning technique with deterministic forecasts as a downscaling method, and as a way to better understand the mechanistic links between oceanographic processes and toxicity.

Acknowledgments

Funding for this work came from the National Science Foundation Research Experience for Undergraduates site at Bigelow Laboratory (award # 1460861), National Atmospheric and Space Administration grants (award NNX16AG59G, EPSCoR award EP-20-01), and Bigelow Laboratory institutional funds.

Literature Cited

Abadi, M., et al. 2015. TensorFlow: large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Allaire, J. J., and F. Chollet. 2018. keras: R Interface to 'Keras'. R package version 2.1.6. https://CRAN.R-project.org/package=keras

Amin, R., J. Zhou, A. Gilerson, B. Gross, F. Moshary, and S. Ahmed. 2009. Novel optical techniques for detecting and classifying toxic dinoflagellate Karenia brevis blooms using satellite imagery. Optics Express 17:9126–9144.

Anderson, D. M., P. Hoagland, Y. Kaoru, and A. W. White. 2000. Estimated annual economic impacts from harmful algal blooms (HABs) in the United States (No. WHOI-2000-11). National Oceanic And Atmospheric Administration Norman Ok National Severe Storms Lab, Woods Hole, Massachusetts, USA.

Anderson, D. M., B. A. Keafer, J. L. Kleinsdinst, D. J. McGillicuddy Jr, J. L. Martin, K. Norton, C. H. Pilskaln, J. L. Smith, C. R. Sherwood, and B. Butman. 2014. *Alexandrium fundyense* cysts in the Gulf of Maine: long-term time series of abundance and distribution, and linkages to past and future blooms. Deep Sea Research Part 2 Topical Studies in Oceanography 103:6–26.

Anderson, C. R., S. K. Moore, M. C. Tomlinson, J. Silke, and C. K. Cusack. 2015. Living with harmful algal blooms in a changing world: strategies for modeling and mitigating their effects in coastal marine ecosystems. Pages 495–561 *in* J. F. Shroder, J. T. Ellis, and D. J. Sherman, editors. Coastal and Marine Hazards, Risks, and Disasters. Elsevier, Amsterdam, The Netherlands.

Anon, A. O. A. C. 2011. Official method 2011.02 Determination of paralytic shellfish poisoning toxins in mussels, clams, oysters and Scallops. Post-column Oxidation Method (PCOX). First Action 2011. AOAC International, Gaithersburg, Maryland, USA.

Apurva, N., A. C. Thomas, and M. E. Borsuk. 2013. Interannual variability in the timing of New England shellfish toxicity and relationships to environmental forcing. Science for the Total Environment 255–266.

Bean, L. L., J. D. McGowan, and J. W. Hurst. 2005. Annual variations of paralytic shellfish poisoning in Maine, USA 1997-2001. Deep-Sea Research II 52:2834–2842.

Burrell, J. 2016. How the machine 'thinks': understanding opacity in machine learning algorithms. Big Data and Society 3:2053951715622512.

Chollet, F., and J. J. Allaire. 2018. Deep Learning with R. Manning Publications Company, Shelter Island, New York, USA.

Chollet, F., et al. 2015. Keras. https://keras.io

Cusick, K. D., and G. S. Sayler. 2013. An overview on the marine neurotoxin, saxitoxin: genetics, molecular targets, methods of detection and ecological functions. Marine Drugs 11:991–1018.

Davidson, K., D. M. Anderson, M. Mateus, B. Reguera, J. Silke, M. Sourisseau, and J. Maguire. 2016. Forecasting the risk of harmful algal blooms. Harmful Algae 53:1–7.

De Silva, S. S., and D. Soto. 2009. Climate change and aquaculture: potential impacts, adaptation and mitigation. Pages 151–212 *in* K. Cochrane, C. De Young, D. Soto, and T. Bahri, editors. Climate change implications for fisheries and aquaculture: overview of current scientific knowledge. FAO Fisheries and Aquaculture Technical Paper. No. 530. FAO, Rome, Italy.

Escalera, L., B. Reguera, T. Moita, Y. Pazos, M. Cerejo, J. M. Cabanas, and M. Ruiz-Villarreal. 2010. Bloom dynamics of *Dinophysis acuta* in an upwelling system: in situ growth versus transport. Harmful Algae 9:312–322.

Guallar, C., M. Delgado, J. Diogene, and M. Fernandez-Tejedor. 2016. Artificial neural network approach to population dynamics of harmful algal blooms in Alfacs Bay (NW Mediterranean): case studies of *Karlodinium* and *Pseudo-nitzschia*. Ecological Modelling 338:37–50.

Hu, C., B. Murch, A. A. Corcoran, L. Zheng, B. B. Barnes, R. H. Weisberg, K. Atwood, and J. M. Lenes. 2016. Developing a smart semantic web with linked data and models for near-real-time monitoring of red tides in the Eastern Gulf of Mexico. IEEE Systems Journal 10:1282–1290.

Kang, H. Y., R. A. Rule, and P. A. Noble. 2011. Artificial Neural Network Modeling of Phytoplankton Blooms and its Application to Sampling Sites within the Same Estuary. Treatise on Estuarine and Coastal Science 161–172.

Kingma, D. P., and J. Ba. 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv: 1412.6980.

Kleindinst, J. L., D. M. Anderson, D. J. McGillicuddy Jr, R. P. Stumpf, K. M. Fisher, D. A. Couture, M. J. Hickey, and C. Nash. 2014. Categorizing the severity of paralytic shellfish poisoning outbreaks in the Gulf of Maine for forecasting and management. Deep Sea Research Part 2 Topical Studies in Oceanography 103:277–287.

Liu, W., Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi. 2017. A survey of deep neural network architectures and their applications. Neurocomputing 234:11–26.

McGillicuddy, D. J., D. M. Anderson, A. R. Solow, and D. W. Townsend. 2005. Interannual variability of

Alexandrium fundyense abundance and shellfish toxicity in the Gulf of Maine. Deep-Sea Research II 52:2843–2855.

McGillicuddy, D. J., D. W. Townsend, R. He, B. A. Keafer, J. L. Kleindinst, Y. Li, J. P. Manning, D. G. Mountain, M. A. Thomas, and D. M. Anderson. 2011. Suppression of the 2010 *Alexandrium fundyense* bloom by changes in physical, biological, and chemical properties of the Gulf of Maine. Limnology and Oceanography 56:2411–2426.

Qin, M., Z. Li, and Z. Du. 2017. Red tide time series forecasting by combining ARIMA and deep belief network. Knowledge-Based Systems 125:39–52.

Record, N. R., B. Tupper, and A. J. Pershing. 2018. The jelly report: forecasting jellyfish using email and social media. Anthropocene Coasts 1:34–43.

Rourke, W. A., C. J. Murphy, G. Pitcher, J. M. van de Riet, B. G. Burns, K. M. Thomas, and M. A. Quilliam. 2008. Rapid postcolumn methodology for determination of paralytic shellfish toxins in shellfish tissue. Journal of AOAC International 91:589–597.

Shumway, S. E., S. Sherman-Caswell, and J. W. Hurst. 1988. Paralytic shellfish poisoning in Maine: monitoring a monster. Journal of Shellfish Research 7:643–652.

Shutler, J. D., K. Davidson, P. I. Miller, S. C. Swan, M. G. Grant, and E. Bresnan. 2012. An adaptive approach to detect high-biomass algal blooms from EO chlorophyll-a data in support of harmful algal bloom monitoring. Remote Sensing Letters 3:101–110.

Simpson, R. H. 1971. A proposed scale for ranking hurricanes by intensity. Minutes of the Eighth NOAA, NWS Hurricane Conference, Miami, Florida, USA.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 15:1929–1958.

Stüken, A., R. J. Orr, R. Kellmann, S. A. Murray, B. A. Neilan, and K. S. Jakobsen. 2011. Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. PLOS ONE 6:e20096.

Stumpf, R. P., M. C. Tomlinson, J. A. Calkins, B. Kirkpatrick, K. Fisher, K. Nierenberg, R. Currier, and T. T. Wynne. 2009. Skill assessment for an operational algal bloom forecast system. Journal of Marine Systems 76:151–161.

Thomas, A. C., A. J. Pershing, K. D. Friedland, J. A. Nye, K. E. Mills, M. A. Alexander, N. R. Record, R. Weatherbee, and M. E. Henderson. 2017. Seasonal trends and phenology shifts in sea surface temperature on the North American northeastern continental shelf. Elementa Science of the Anthropocene 5:48.

Van De Riet, J. M., R. S. Gibbs, F. W. Chou, P. M. Muggah, W. A. Rourke, G. Burns, K. Thomas, and M. A. Quilliam. 2009. Liquid chromatographic postcolumn oxidation method for analysis of paralytic shellfish toxins in mussels, clams, scallops, and oysters: single-laboratory validation. Journal of AOAC International 92:1690–1704.

Velo-Suárez, L., and J. C. Gutiérrez-Estrada. 2007. Artificial neural network approaches to one-step weekly prediction of *Dinophysis acuminata* blooms in Huelva (Western Andalucía, Spain). Harmful Algae 6:361–371.

Wang, L., et al. 2019. An approach of improved Multivariate Timing-Random Deep Belief Net modelling for algal bloom prediction. Biosystems Engineering 177:130–138.

Yentsch, C. M., C. M. Lewis, and C. S. Yentsch. 1980. Biological resting in the Dinoflagellate *Gonyaulax excavata*. BioScience 30:251–254.