Version of Record: https://www.sciencedirect.com/science/article/pii/S0380133020301817
Manuscript_25fdeb26a0f0c4e19227872951f90a28

Using Machine Learning to Model and Predict Water Clarity in the Great Lakes

Cameron C. Lee*[a]

Brian B. Barnes[b]

Scott C. Sheridan[a]

Erik T. Smith[a]

Chuanmin Hu[b]

Douglas E. Pirhalla[c]

Varis Ransibrahmanakul[c]

Ryan Adams[a]

[a] Kent State University; Department of Geography; 413 McGilvrey Hall; 325 South Lincoln Street; Kent, Ohio 44242, USA

[b] University of South Florida; College of Marine Science; 140 7th Avenue South, St. Petersburg, FL 33701, USA

[c] National Oceanic and Atmospheric Administration; National Centers for Coastal Ocean Science; 1305 East West Highway, Rm 8110; Silver Spring, MD 20910, USA

*** Corresponding Author Additional Information:**

Phone: +1 330-672-0360

Fax: +1 330-672-4304

Email: cclee@kent.edu

ABSTRACT:

Over the last several decades, multiple environmental issues have led to dramatic changes in the water clarity of the Great Lakes. While many of the key factors are well-known and have direct anthropogenic origins, climatic variability and change can also impact water clarity at various temporal scales, but their influence is less often studied. Building upon a recent examination of the univariate relationships between synoptic-scale weather patterns and water clarity, this research utilizes nonlinear autoregressive models with exogenous input (NARX models) to explore the multivariate climate-to-water clarity relationship. Models trained on the observation period (1997-2016) are extrapolated back to 1979 to reconstruct a daily-scale historical water clarity dataset, and used in a reforecast mode to estimate real-time forecast skill. Of the 20 regions examined, models perform best in Lakes Michigan and Huron, especially in spring and summer. The NARX models perform better than a simple persistence model and a seasonal-trend model in nearly all regions, indicating that climate variability is a contributing factor to fluctuations in water clarity. Further, six of the 20 regions also show promise of useful forecasts to at least 1 week of lead-time, with three of those regions showing skill out to two months of lead time.

KEYWORDS: Great Lakes, water clarity, climate variability, machine learning, synoptic climatology, climate change

INTRODUCTION

The Great Lakes contain 84% of North America's fresh water and provide critical resources to the local economy and environment (e.g., Great Lakes related jobs accounted for $54 billion in annual compensation in 2007 in the state of Michigan alone; Vaccaro et al., 2009), with over 30% of the gross domestic product of the United States and Canada coming from the region (Krantzberg and Boer, 2006, Smith et al., 2020). The Great Lakes provide countless services and opportunities for millions of Americans, supporting tourism and other economic uses such as fishing, boating, beach use and scuba diving. However, over the past decade, changes in human use and the introduction and migration of non-native mussel populations throughout the Great Lakes have rapidly altered the system (Nalepa et al., 2009), leading to changes in various water quality metrics, including water clarity. Due partly to this increase in the mussel population, many of the Great Lakes have experienced increasing water clarity (clearer water) via mussel filtering of phytoplankton (Binding et al., 2015, Cha et al., 2013). In turn, the reduced phytoplankton has diminished primary production, reducing the availability of food for fish and generally disrupting the natural ecosystem. The increased water clarity has also enhanced light penetration to the lake bottom, stimulating benthic plant growth (Ricciardi et al., 1997; Skubinna et al., 1995). At the same time, especially in Lake Erie, agricultural practices over the past several decades have led to decreased water clarity due to nutrient runoff from phosphorus-based fertilizers (Jarvie et al., 2017), resulting in seasonal algal blooms, some of which have contaminated city water systems (Steffen et al., 2017).

These changes to Great Lakes water clarity stem mostly from anthropogenic factors. However, climate change and shorter-term climate variability can also influence water clarity, though these factors are less well-understood. Precipitation has been shown to impact water clarity via suspended particles and organic matter discharged into the Great Lakes from connected streams, through water temperatures, wind speeds and other meteorological variables are also known to influence algal blooms and thus, water clarity (Michalak et al., 2013; Wynne et al., 2010; Smith et al., 2020). In the longer-term, rising temperatures and shifts in precipitation patterns and runoff will likely alter water clarity via changes in algal blooms (Paerl and Paul, 2012). Additionally, compounding factors including decreased ice extent, lake temperatures that are increasing at a faster rate than atmospheric warming (Austin and Colman, 2007; Woolway and Merchant, 2018), longer growing seasons, and altered precipitation/discharge patterns have become more problematic during the last century. Such factors can cause further declines in water quality, as well as negative impacts on human health and infrastructure.

In a recent study, Smith et al. (2020) described the development of a water clarity index (KDI) for the Great Lakes and elucidated some of the relationships between synoptic-scale meteorological patterns and water clarity in the region. While previous studies have examined water clarity variability and trends in the Laurentian Great Lakes (e.g. Binding et al., 2015; Yousef et al., 2017; Dobiesz and Lester, 2009; Auer et al., 2010) or modeled changes in other lake parameters (e.g. lake temperatures in Trumpickas et al., 2009), to the authors' knowledge none have attempted to model water clarity largely based upon historical meteorological forcings. Thus, building upon Smith et al., (2020), this research describes the development of a series of artificial neural network (ANN)-based time-series models to empirically predict shorter-term variability and longer-term trends in water clarity in the Great Lakes. Once validated, these ANN models are then simulated on historical climate data to generate a complete reconstructed time series of water clarity in the Great Lakes from 1979 to present, and applied to forecast model data to produce a 3 year reforecast of water clarity, and an ongoing real-time water clarity outlook.

METHODS

While we provide an overview of the data acquisition and treatment, the regionalization, and the process of atmospheric pattern classification used in this research, further detail can be found in Smith et al. (2020).

Data and Treatment

Water clarity data were derived as described in Smith et al. (2020). Briefly, all SeaWiFS (Sea-viewing Wide Field-of-view Sensor onboard the satellite Orbview-2; 1997-2010; ~1.1km spatial resolution at nadir), MODIS/A (Moderate Resolution Imaging Spectroradiometer onboard Aqua; 2002-present; ~1km spatial resolution at nadir), and VIIRS/SNPP (Visible Infrared Imaging Radiometer Suite onboard Suomi-NPP; 2012-present; ~750m spatial resolution at nadir) remote sensing reflectance (R_{rs}; sr^{-1}) data covering the Great Lakes region were downloaded at Level-2 from NASA's Ocean Biology Distributed Active Archive Center (OB.DAAC). Data for all three sensors correspond to NASA processing version R2018.0, which were mapped to an equidistant cylindrical projection with spatial resolution of 1km and geographic bounds of 41° to 49°N, 76° to 92°W. The Kd_lee algorithm (Lee et al 2005) was used to derive K_d(488) (diffuse attenuation coefficient for downwelling light at 488 nm) from these R_{rs} data. Log-transformed SeaWiFS and VIIRS K_d(488) data were scaled to match MODIS using linear regression coefficients determined from collocated and coincident satellite / satellite matchups. Specifically, y=10^($\beta_1$log(x) + β_0), where x is the original SeaWiFS or VIIRS/SNPP K_d(488) data, y is the scaled value (to match MODIS/A), and the coefficients β_0 and β_1 are 0.0069477 and 1.0087 (-0.045826 and 1.1369) for SeaWiFS (VIIRS/SNPP). KDI were then calculated for each 1km pixel as the normalized anomaly in K_d(488) at various temporal resolutions. For example, monthly resolution MODIS data were used in regionalization analyses, whereby:

$KDI_{monthly}$=[monthly_mean_K_d(488) - monthly_mean_K_d(488)_climatology] / monthy_stdev_K_d(488)_climatology

In order to analyze the multivariate climate-to-KDI relationship, the study domain was separated into 20 different regions based upon the $KDI_{monthly}$ data for each of the 1km pixels. As described in Smith et al. (2020), multiple clustering techniques and multiple cluster numbers were evaluated using 5 different cluster validation metrics. Among the nearly 50 options tested, a k-means based 15 cluster/region solution was ultimately selected. While the majority of the 15 regions were geographically cohesive, some were split across multiple lakes. Thus, the final

20 region solution shown in Figure 1 (and used for all further analyses) was based upon a re-drawing of these regions in ArcMap, using the original 15-cluster solution as a framework.

After regionalization, KDI_{daily} datasets were calculated similarly to $KDI_{monthly}$, comparing daily mean $K_d(488)$ values to climatological values defined as 30-day running mean and standard deviation. Prior to KDI_{daily} calculation, a 5x5 median filter was applied to the daily mean $K_d(488)$ data. Mean KDI_{daily} within each of 20 regions were calculated for all dates with at least 5% valid pixel cover (see Smith et al., 2020), and this 'regional mean KDI_{daily}' is used in all further analyses (and is referred to simply as KDI hereafter).

Stream discharge (m^3/s) data were gathered from the U.S. (U.S. Geological Survey, 2017) and Canadian (Water Survey of Canada, 2017) streamflow gauging stations near each lake. Only stations with at least 25 years of data since 1979 were used. Gauging stations were divided according to the lake regions in which the stream ends; however, the nearest station from an adjacent region was used for regions which had no gauging station with at least 25 years of data.

All historical climate data were obtained from the North American Regional Reanalysis (NARR; Mesinger et al., 2006) at the native 32km spatial resolution. Daily mean fields of sea-level pressure (SLP), 500mb geopotential heights (500Z), 850mb temperatures (850T) and 10-m zonal and meridional wind components were obtained for 1979-2015 for the spatial domains shown in Figure 1. Because the daily pressure and geopotential height gradient was determined to be more important in forcing regional KDI variability, raw 500Z and SLP data were normalized by subtracting the daily mean value from each grid point (i.e. creating spatial anomalies/gradients), which also helped reduce the seasonality of these variables. These same atmospheric circulation variables were also retrieved from the Climate Forecast System (CFS) for reforecasting (CFS-RF; January 2016 to December 2018) of the KDI. As CFS data are output on a different grid (0.5° x 0.5° resolution) than NARR, all CFS data were spatially interpolated to the NARR grid by using a Delaunay triangulation of the scattered two-dimensional data points.

Delaunay triangulation tends to be more efficient and suffers from fewer artifacts than other interpolation methods, such as inverse distance weighting (Amidror, 2002).

Prior to classification, all NARR atmospheric variables (raw fields of 850T and 10-m wind, and spatial anomaly fields of SLP and 500Z) were standardized for each location (i.e. z-scored by column), and then subjected to an s-mode principal components analysis (PCA – where rows are days and columns are locations; Yarnal, 1993). Principal component scores (PCs) with eigenvalues greater than 1 were retained and used in classification. An artificial neural network (ANN) clustering procedure known as self-organizing maps (SOMs) were then used to classify each set of PCs into a set of discrete atmospheric patterns (e.g. Figure 2 for 500Z; see Electronic Supplementary Material (ESM) Figs. S1-S3 for the other SOMs), with the ideal SOM architecture for each dataset selected based upon multiple cluster validation metrics (described in Smith et al., 2020). In order to create CPs (climate parameters) from the CFS dataset that corresponded to the CPs originally developed from the NARR dataset, CFS data treatment must follow that of NARR data. Thus, spatial gradients of 500Z and SLP CFS data were computed and then all variables were standardized and turned into PCs using a two-step process. First, 'virtual CFS z-scores' were calculated by subtracting the NARR gridpoint means from the interpolated CFS data and dividing by the NARR standard deviations (at each interpolated gridpoint). These z-scores were then multiplied by the loadings matrix derived from the NARR-based PCA, creating 'virtual CFS PC scores' to input into the SOM neural network. The SOM is then run (not re-trained) on the retained virtual PCs, producing a daily-scale classification of CFS data.

In addition to the CP classifications, which assess the atmospheric circulation over broad areas, local surface weather conditions will also be evaluated for their link to the KDI. These conditions were assessed via the Gridded Weather Typing Classification (GWTC; Lee, 2015). The GWTC uses of eight-times daily values of 6 different variables (2-m temperature, 2-m dew point, 10-m wind speed, 10-m wind direction, SLP, and total cloud cover) from the NARR dataset (at ~60km resolution) to identify 11 different spatiotemporally relative weather types (WTs; Lee, 2015), similar to traditional air masses. Calendars of daily historical and real-time forecast WT data from 1979-2018 were obtained from the GWTC homepage (GWTC, 2020) for areas over the Great Lakes. Since some of the 20 regions had multiple GWTC locations within their borders, a regional daily WT was derived by calculating the mode WT for all GWTC locations within each region's boundaries.

NARX Modeling

Modeling of the climate-KDI relationship was completed using nonlinear autoregressive models with exogenous input (NARX models), an ANN-based time-series modeling framework that accounts for both the nonlinear relationships between predictors (climate variables) and predictand (KDI), while also incorporating the natural autocorrelation of both (Lee et al., 2017). However, to help minimize the computational time and eliminate collinear variables from entering the model, input variable selection (IVS) must first be completed. The possible set of predictor variables included 5 categorical variables: the 4 sets of SOMs, and the region's GWTC WT (which were all turned into dummy variables), along with either the 4 (850T) or 5 (SLP, Wind and 500z) leading PCs from the circulation pattern classification process described above, up to 4 stream discharge gauges for each region (that showed significant correlations with KDI), a sinusoidal seasonal signal, a linear trend line, and lake temperatures. Each of these variables was then transformed into a lagged matrix of 0-28 days, based upon the maximum amount of significant lag between each variable and KDI in prior testing. These lagged matrices were then standardized and subjected to a PCA, producing a set of orthogonal PCs that have now accounted for both autocorrelation and collinearity in the set of potential predictor variables. Spearman correlations between each of these PCs and the regional KDI time series were then computed, with all significantly correlated ($p<0.001$) PCs retained for input into NARX modeling.

As with all machine learning methods, NARX models learn relationships through iterative training. First, the time series of predictors (significant PCs from IVS) and predictands (regional KDI) are separated into three cohesive time-blocks: training (80%), internal validation (10%) and testing (10%). The modeler must also set two parameters, the number of neurons in the ANN – which equates to the amount of complexity and interaction in the model; and the number of delays/lags to include in the model. Using Levenberg-Marquardt optimization, through multiple iterations, the model then adjusts the value of weights (which connect each predictor to each neuron) and biases of each neuron, in an effort to minimize the mean squared error (MSE) on the training set of data. Using a technique known as early-stopping, after each of these adjustments, the updated model is then run on the internal validation block of data and the MSE is also computed. If, after a user-defined number (5) of steps, the MSE on the internal validation block of data fails to improve (decrease), then the model trained 5 steps prior is considered optimized. While early-stopping considerably cuts computational time, it also helps mitigate overfitting, as the final model can be considered optimized to two separate portions of the dataset. However, because the internal validation block of data is not completely independent of model training (as it determines when the model stops training), the testing block of data is held out for external validation of model performance.

Due to natural autocorrelation, the best predictor of most environmental variables (including KDI) is its value from the prior time step (the lag-1 value). Accordingly, the training initially proceeds using an 'open-loop' (OL) framework, whereby the model is fed the previous day's values of observed KDI as an input into predicting the next day's output. However, for *real-time* prediction for multiple time-steps ahead, there are no actual observed KDI values yet. Thus, once optimized on open-loop, each NARX model is then trained and run in a 'closed-loop' (CL) framework, whereby the model proceeds chronologically, feeding its own model output back into itself as a predictor at the next time step.

Three other considerations must also be made while training NARX models. First, the ideal model parameters (i.e. the number of neurons and delays) are unknown and not easily estimated, though they should both be minimized to curtail computational time and the possibility of overfitting. To overcome this hurdle, a brute-force technique was used where every possible combination of the number of neurons (from 1 to 10) and delays (from 1 to 5) were used as settings in separate NARX models which modeled 50 different possible model-architectures.

Secondly, the weight and bias terms are initialized with random values, meaning that even with all other model settings being the same, each model will produce a slightly different result. Thus, every NARX model was trained 10 times and ensemble medians of these 10 permutations were used as the final model output time series, including those trained (and discarded) in brute-force testing of model-architectures.

Finally, while the testing block of data is independent, this leaves the modeler with only 10% of the dataset from which to draw conclusions about model performance. However, as described by Lee et al., (2017), multiple NARX models, each with a separate 10% chunk of data designated as the testing block of data (and correspondingly different blocks of data then designated as training and internal validation), can be trained, and then the testing portions of each of these models can be reconstructed to form a complete time series of independent output data. Further, in this research, with the same block of data as the training block, the internal validation and independent testing blocks can be flipped, meaning each 10% block of data can be used twice as the testing block, with slightly different portions of the models being used for training for each, resulting in 20 different possible 'settings' for dividing the data into time blocks. These reconstructed independent testing ensemble (RITE) datasets are used below for calculating OL and CL model performance statistics for the historical period (hindcast; 1979-2015).

Considering these 20 settings, the 10 permutations for constructing ensembles, and the 50 different possible model-architectures, 10,000 NARX models were trained for each of the 20 regions – or 200,000 total models. For each region, among the 50 possible options, the winning set of 200 NARX models (20 settings x 10 permutations of the best model-architecture) was ultimately chosen, and saved for use with CFS-RF data for reforecasting KDI.

There are three different ways in which the models can be run: 1) OL hindcast, run once from the beginning of the time series to the end on open-loop (as described above); 2) CL hindcast, run once from the beginning to the end on closed-loop (as described above), and 3) CL reforecast, run in a 'reforecast mode' on closed-loop, which means the models are run for each day in the time series (as if one were making a daily forecast in real-time) and the actual observed past values of KDI are input to initialize the day's model before it runs forward in closed-loop mode to produce multi-step ahead forecasts. Each of these modes of operation will produce statistics that will provide evidence of the performance of different aspects of the models: 1) OL hindcast performance which gives us an idea of how well the process can be modeled in general, and of the potential lead-1 forecast performance in real-time; 2) the CL hindcast performance gives us an idea of the confidence we can have in the 'accuracy' of our hindcast; and 3) a CL reforecast performance gives us an idea of what the real-time forecast performance will be within the lead-time horizons of interest.

Overall NARX model performance was evaluated using Spearman correlations between the actual observed KDI time series, hit rates for extreme KDI events (both extreme clear-water and extreme cloudy-water events, demarcated as the top and bottom 20th percentiles of the time series), and the improvement of the NARX model over simpler models (e.g. that of a seasonal cycle and trend model, or a simple persistence model). These metrics are stratified in several ways: correlation of the whole time series, monthly averaged correlation (to estimate seasonal model skill), and annually averaged correlations (to examine interannual variability of model performance).

RESULTS AND DISCUSSION

Model performance varies by region and season, and expectedly, is much better using the OL framework (Table 1). Generally, the regions and seasons with the best OL model performance also had substantially better CL model performances. Seasonally, for most regions, the NARX models performed poorest in autumn and best in spring, followed closely by summer (Table 2). Geographically, NARX models are able to best recreate the daily KDI time series in Lakes Huron and Michigan – Regions 6, 8, 9 (in Huron) and Regions 17, 18, 19, 20 (in Michigan) each had open-loop rho>0.84 and closed-loop rho>0.65, with region 20 performing best (OL rho=0.92, CL rho=0.79). These qualitatively high correlations on OL highlight not only the flexibility of NARX models in being able to 'learn' the climate-KDI relationship, but also the potential for real-time forecasting. These better modeled regions share some similar characteristics in their time series (Figure 3), in which the early observational period (1997-2003) is dominated by large seasonal variability and slightly decreasing KDI values (clearing water), followed by an abrupt drop in KDI around 2004, and then a much more inconsistent seasonality and slow decline in KDI thereafter. These general trends in the water clarity of the Lake Huron and Michigan regions are consistent with results found in Binding et al (2015) and others who attribute much of this to reductions in phytoplankton biomass due to the colonization of invasive quagga mussels which increase water clarity via filter-feeding and also dramatically reduce spring blooms. Decreased phosphorous loading has also been suggested as a possible cause for these trends (Yousef et al., 2017).

Relative to a simple persistence model, where the 'predicted KDI' is equal to the previous day's KDI, OL models show substantially better skill (Table 1: R2%ImpAutoC = improvement of the r-squared over the lag-1 persistence). For example, in Region 7, 1-day autocorrelation of KDI is only r=0.40, however, OL performance is markedly better (rho=0.74, r=0.80), yielding 48% more variability explained with this model. The better performing regions with this metric are those with the poorest CL model performance (especially relative to OL model performance). This is unsurprising considering that the main difference between OL and CL is that the former has actual observed KDI values as input, while the latter relies more on the exogenous (climate) variables. Still, this metric elucidates the relative importance of climate in forcing water clarity, which is mostly dominated by autocorrelation in the short-term, and other anthropogenic factors over the longer term (e.g. invasive mussels, fertilizer use). This also highlights the potential value-added in real-time forecasting of water clarity with climate variables. The ability of the NARX models to predict extreme KDI events largely coincides with the correlative analyses – Lakes Huron and Michigan have the best hit rates, with Regions 9 and Region 20 having hit rates greater than 70% for high-KDI events, and greater than 53% for low-KDI events. Generally, extreme high-KDI events are more skillfully predicted than the low-KDI events.

Compared to a model that just uses the seasonal cycle and a quadratic trend to 'predict' KDI, CL models do show a slight improvement of the r-squared in most regions, with some of the largest improvements in Lake Huron, especially R8 (22% improvement) and R9 (12% improvement). Overall, however, these relatively low scores indicate that, in CL hindcast mode, a large part of the model skill is derived more from seasonal cycles and/or trends in KDI more-so than from the day-to-day weather variability. Important to note is that other sets of NARX models were fully trained and run on the *anomalies* derived from these seasonal/trend curves, but yielded negligibly different performance results, which underscores the robustness of the NARX methodology to the inputs used and its ability to 'learn' the seasonal and long term relationships between climate and KDI.

The aforementioned statistics reflect model performance across the entire time series; however, interannual variability does exist (Table 3). When year-over-year correlations are examined on the testing block of the OL hindcast output, all regions show potential for reliable 1-day lead-time forecasts, with small standard deviations from mean performance, and the best modeled regions having rho>0.5 for all years, with some years as high as rho>0.9. However, when examining the yearly output of the CL hindcast data, the skillfulness is much more sporadic. While some regions show rho>0.80 for some years, all but 3 regions (R10, R17, and R20) have at least one year which exhibits a *negative* correlation between NARX model output and actual observations, highlighting the need for caution when interpreting the results of the CL hindcast. Temporally, the NARX models generally perform better during the beginning (1998-2004) and the end of the time series (2010-2015). The poorer-performing middle time period (2004-2010) is expected as it represents the period with the most abrupt (and non-climate-related) changes in water clarity trends and seasonality (especially in Lakes Huron and Michigan) as invasive species began colonizing the Great Lakes.

One of the overarching goals of the grant funding this research was to examine the potential of water clarity to serve as an indicator of climate change. Overall, statistically significant decreases in KDI over the 1979-2015 time period are present in 14 of the 20 regions, along with 2 significantly increasing regions (Region 3 and Region 14; Table 1). While these trends towards clearer water are obvious, teasing out a *climate change* signal from other (largely anthropogenic) non-climate related factors is difficult, a result noted in other research examining changes in fish abundance in the Great Lakes (Wuebbles et al., 2019). Further, with warming waters, one probable result from climate change alone would be towards cloudier water conditions (i.e. increased KDI) accompanying increased rates of algae blooms (e.g. Paerl and Paul, 2012), rather than the general clearing trend that we observed.

While previous research has noted the difficulty in determining the relative importance for each predictor variable that is used in an ANN model, a few different ways have been proposed, including the predictor-constant replacement method (Lee et al., 2017), the random permutation replacement method (Giam and Olden, 2015) and Garson's algorithm (Garson, 1991), among others. The current research explored each of these referenced options, with them all showing similar results. Thus, in keeping with the methodology used by Lee et al., (2017), the predictor-constant replacement method was used, whereby each model member was re-run setting a different input variable to a constant (zeros, the mean of the PCs) and the change in model performance (using median absolute errors) was noted. These changes in performance were then multiplied by the retained PC loadings saved from the IVS procedure described above to determine each variable's contribution to the model. Results show that the linear trend included in the model was the most important, especially in the well-modeled regions of Lake Huron and Lake Michigan; and the seasonal signal played the largest role in Lake Huron's sub-regions. However, among the climate-related variables, on average, the PCs of 500z play the largest role, followed by the MSLP PCs, 10m-wind PCs and then discharge (which plays a major role in the better performing regions, e.g. R20 and R9; Table 4). While the categorical CPs play much more minor roles everywhere, the patterns of 850T and 500Z are consistently more important than MSLP and Wind CPs. This is likely due to the strong seasonal variability in the frequency of 850T and 500Z patterns (compared to the MSLP and Wind classifications), as KDI is also highly seasonal. This result, along with the relative importance of stream discharge in impacting water clarity, was noted in previous research using correlative analyses to examine the relationship between KDI and these same SOM-based circulation patterns (Smith et al., 2020). Because KDI is also strongly seasonal (Figure 3), summer-dominant CPs (often relating to more zonal flow and less precipitation) are associated with lower KDI (clearer water), while the winter and spring-dominant CPs, are more often favorable for precipitation events, and are thus associated with higher KDI (Smith et al., 2020). Smith et al. (2020) also noted that Wind and SLP patterns generally correlate more strongly to KDI in Lakes Erie and Ontario, a result which is also noted here. However, these patterns also are important factors in the NARX modeling of parts of Lake Superior, an area where Smith et al. (2020) showed less association. Because SLP and wind play the smallest roles in impacting KDI (perhaps due to their weaker seasonality compared to 500Z and 850T patterns) interpretation of their relative importance from region-to-region is tenuous, especially in Lakes Superior, Erie and Ontario, where NARX model performance is generally poorer.

The best indication of real-time forecast skill at longer lead times can be determined from the performance of the models when run in ‘reforecast mode’ (Figure 4). Herein, we ran the models in reforecast mode on an entirely separate series of data, from 2016-2018. Theoretically, the OL results noted above will be identical to the reforecast model skill at a lead time of 1 time step (i.e. 1 day, in this research) because the default for OL models is to perform 1-step ahead predictions. However, because we convert the OL models to CL and then retrain them (to improve CL performance) and we are using an entirely separate time series of data for reforecasting (2016-2018), our performance in reforecast mode will differ than the OL performances noted above. The effect of this is that at shorter lead-times model performance suffers slightly by comparison to OL results noted above, however, performance improves at longer time horizons relative to persistence. For example, in Region 6, while the lead-1 correlation (r=0.73) between NARX models and observed KDI is only marginally better than the lead-1 persistence model (r=0.70), at 60 days of lead time, the NARX models are considerably better (r=0.58) than 60-day persistence (r=0.12). In fact, all regions show a positive linear trend (from 1- to 60-days of lead time) in the r-squared improvement over persistence in these reforecasts, with the largest improvements generally in the same (better performing) regions of Lake Huron and Lake Michigan noted above. This noted, as with any prediction model, the absolute model skill (i.e. not relative to persistence) for all regions gets worse at longer lead-times, and outside of Lakes Huron and Michigan, most regions' model performance deteriorates to the point of being unhelpful. Still, with correlations approaching r>0.5 for these regions, sub-seasonal forecasts of water clarity may prove valuable.

CONCLUSIONS

This research utilized nonlinear autoregressive models with exogenous input (NARX models) – an ANN-based time-series model – to examine the relationship between meteorological variables and a water clarity index (KDI) in the Great Lakes. NARX models were also used to hindcast KDI from 1979-2015 and to predict KDI over the 2016-2018 time period in a reforecast mode. Results show that regional climate variability is inherent to fluctuations in water clarity. Model performance varied seasonally and geographically, with models for parts of Lake Huron and Lake Michigan performing best, especially in the spring and summer months. Models in nearly all regions show skill beyond that of a simple persistence model (on open-loop) or a seasonal-trend model (on closed-loop), which highlights the usefulness of the modeling methodology and the relative importance of integrating meteorological variables into models of water clarity in these lakes. Of the 20 regions examined, models trained in six of the regions show promise for real-time forecasting past 1 week of lead time (r>0.5), and three of those regions at up to 60 days of lead-time.

While this research shows promise for real-time predictions of Great Lakes water clarity using NARX models, there are a number of limitations that must be considered when interpreting the results herein. First, all of the results above are essentially based upon 'perfect' forecasts, i.e the exogenous (climate) variables, even when models are run in 'reforecast mode', are the actual *observed* (from reanalysis) weather conditions within each 60-day lead-time window, rather than the *forecasted* conditions for those days. Thus, due to deteriorating skill of weather/climate model forecasts at longer lead times, the results noted herein will likely suffer when applied in real-time. Still, with reliable weather prediction out to about 1 week, and the relative dependence of modeled KDI on previous values of itself (rather than climate models), its trends and its seasonal signal, the impact of this limitation on real-time performance may well be minimal, but undoubtedly deserves future attention.

A second limitation is the lack of non-climate related variables as predictors of water clarity. While it is well documented that increased fertilizer use, invasive mussels, land use patterns and soil types within specific watersheds and a host of other anthropogenic and natural phenomena can impact water clarity in the Great Lakes, adequate data for these factors, at the necessary spatial and temporal scales for our modeling purposes, is non-existent. For this same reason, an input as important as discharge data was even relegated to a seasonal cycle, since real-time forecast data out to 60 days of lead time is not available for the streams found to have the greatest impact on KDI in each of our regions. As a result, there is a substantial amount of unexplained variability in KDI that simply cannot be modeled (short of collecting such data sets somehow in the future), and thus, despite the potential for real-time forecasting, the hindcasts of KDI developed herein should be used with caution, especially those outside of Lakes Huron and Michigan.

In terms of potential future directions stemming from this research, currently ongoing research has begun real-time forecasting (or a Great Lakes water clarity outlook) in order to build up a long enough period of record for analyzing model performance in this capacity, while another future research avenue could use actual reforecasts archived from major modeling centers, such as NOAA's Climate Forecast System Reforecast to examine the same. Due to the importance of discharge in forcing KDI, planned research will be looking to improve the skill of KDI forecasts by adapting the NARX modeling framework to model/forecast discharge at the lead times of interest (e.g. 60 days) based upon climatological forcing, and then nesting these discharge models inside the KDI models. Future case studies of the impact of specific extreme precipitation events on KDI may also help shed light on the contribution of stream discharge of water clarity variability (e.g. Cooney et al., 2018). Incorporating lake water-level data, which may partially govern rates of coastal erosion and sediment resuspension via lake bed erosion in shallower waters (Dusini et al., 2009; Valipour et al., 2017), may also enhance NARX model accuracy in future research.

In addition to providing new information on sensitive regions of the Great Lakes subjected to fluctuations and trends in water clarity, a major goal of this research was to enhance awareness of the long-term changes and variations taking place in water clarity across the entire Great Lakes system. Particularly relevant are water clarity estimates near natural and cultural resources found in the Great Lakes that possess exceptional historic, archaeological, and recreational value to the region. Future research efforts focused on real-time satellite monitoring and forecasting of water clarity conditions will likely lead to more coordinated and timely management response to water clarity events, and improved recreational activity planning for users such as divers, kayakers, and snorkelers, thereby increasing regional tourism and facilitation of public access through NOAA, the National Park Service and other protected area management agencies in the region.

ACKNOWLEDGEMENTS

This research was funded by the National Aeronautics and Space Administration's (NASA's) Research Opportunities in Space and Environmental Sciences (ROSES) funding opportunity, Development and Testing of Potential Indicators for the National Climate Assessment, Award NNX16AH12G. We thank NASA's Ocean Biology Processing Group (OBPG) for providing MODIS data (https://oceancolor.gsfc.nasa.gov). We would also like to thank the editor and the anonymous reviewers for their comments.

REFERENCES

Amidror, I. (2002). Scattered data interpolation methods for electronic imaging systems: a survey. Journal of electronic imaging, 11(ARTICLE), 157-76.

Auer, M.T., Tomlinson, L.M., Higgins, S.N., Malkin, S.Y., Howell, E.T. and Bootsma, H.A., 2010. Great Lakes Cladophora in the 21st century: same algae—different ecosystem. Journal of Great Lakes Research, 36(2), pp.248-255.

Austin, J.A. and Colman, S.M., 2007. Lake Superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback. Geophysical Research Letters, 34(6) https://doi.org/10.1029/2006GL029021.

Binding, C.E., Greenberg, T.A., Watson, S.B., Rastin, S. and Gould, J., 2015. Long term water clarity changes in N orth A merica's G reat L akes from multi-sensor satellite observations. Limnology and Oceanography, 60(6), pp.1976-1995.

Cha, Y., Stow, C.A. and Bernhardt, E.S., 2013. Impacts of dreissenid mussel invasions on chlorophyll and total phosphorus in 25 lakes in the USA. Freshwater Biology, 58(1), pp.192-206.

Cooney, E.M., McKinney, P., Sterner, R., Small, G.E. and Minor, E.C., 2018. Tale of two storms: Impact of extreme rain events on the biogeochemistry of Lake Superior. Journal of Geophysical Research: Biogeosciences, 123(5), pp.1719-1731.

Dobiesz, N.E. and Lester, N.P., 2009. Changes in mid-summer water temperature and clarity across the Great Lakes between 1968 and 2002. Journal of Great Lakes Research, 35(3), pp.371-384.

Dusini, D.S., Foster, D.L., Shore, J.A. and Merry, C., 2009. The effect of Lake Erie water level variations on sediment resuspension. Journal of Great Lakes Research, 35(1), pp.1-12.

Garson, D.G., 1991. Interpreting neural network connection weights. AI Expert 6(4), 47-51.

Giam, X. and Olden, J.D., 2015. A new R2-based metric to shed greater insight on variable importance in artificial neural networks. Ecological Modelling, 313, pp.307-313.

GWTC, 2020. The Gridded Weather Typing Classification (GWTC) Data page. Accessed at: http:// www.personal.kent.edu/ ~cclee/gwtcdata.html on 10-Feb-2020.

Jarvie, H.P., Johnson, L.T., Sharpley, A.N., Smith, D.R., Baker, D.B., Bruulsema, T.W. and Confesor, R., 2017. Increased soluble phosphorus loads to Lake Erie: Unintended consequences of conservation practices? *Journal of environmental quality*, *46*(1), pp.123-132.

Krantzberg, G. and Boer, C.D., 2006. A valuation of ecological services in the Great Lakes basin ecosystem to sustain healthy communities and a dynamic economy. Prepared for the Ontario Ministry of Natural Resources by Dofasco Centre for Engineering and Public, Policy McMaster University. Retrieved Mar, 8, p.2007.

Lee, C.C., 2015. The development of a gridded weather typing classification scheme. International Journal of Climatology, 35(5), pp.641-659.

Lee, C.C., Sheridan, S.C., Barnes, B.B., Hu, C., Pirhalla, D.E., Ransibrahmanakul, V. and Shein, K., 2017. The development of a non-linear autoregressive model with exogenous input (NARX) to model climate-water clarity relationships: reconstructing a historical water clarity index for the coastal waters of the southeastern USA. Theoretical and Applied Climatology, 130(1-2), pp.557-569.

Lee, Z., K-P Du, and R. Arnone, 2005: A model for the diffuse attenuation coefficient of downwelling irradiance. JGR-Oceans. 110(C2). https://doi.org/10.1029/2004JC002275

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H. and Ek, M.B., 2006. North American regional reanalysis. Bulletin of the American Meteorological Society, 87(3), pp.343-360.

Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., Chaffin, J.D., Cho, K., Confesor, R., Daloğlu, I. and DePinto, J.V., 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proceedings of the National Academy of Sciences, 110(16), pp.6448-6452.

Nalepa, T.F., Fanslow, D.L. and Lang, G.A., 2009. Transformation of the offshore benthic community in Lake Michigan: recent shift from the native amphipod *Diporeia* spp. to the invasive mussel *Dreissena rostriformis bugensis*. Freshwater Biology, 54(3), pp.466-479.

Paerl, H.W. and Paul, V.J., 2012. Climate change: links to global expansion of harmful cyanobacteria. *Water research*, *46*(5), pp.1349-1363.

Ricciardi, A., Whoriskey, F.G. and Rasmussen, J.B., 1997. The role of the zebra mussel (*Dreissena polymorpha*) in structuring macroinvertebrate communities on hard substrata. Canadian journal of fisheries and aquatic sciences, 54(11), pp.2596-2608.

Skubinna, J.P., Coon, T.G. and Batterson, T.R., 1995. Increased abundance and depth of submersed macrophytes in response to decreased turbidity in Saginaw Bay, Lake Huron. Journal of Great Lakes Research, 21(4), pp.476-488.

Smith, E.T., Lee, C.C., Barnes, B.B., Adams, R.E., Pirhalla, D.E., Ransibrahmanakul, V., Hu, C., Sheridan, S.C., 2020. A synoptic climatological analysis of the atmospheric drivers of water clarity variability in the Great Lakes. Journal of Applied Meteorology and Climatology. 59(5) 9150935. DOI: 10.1175/JAMC-D-19-0156.1

Steffen, M.M., Davis, T.W., McKay, R.M.L., Bullerjahn, G.S., Krausfeldt, L.E., Stough, J.M., Neitzey, M.L., Gilbert, N.E., Boyer, G.L., Johengen, T.H. and Gossiaux, D.C., 2017. Ecophysiological examination of the Lake Erie Microcystis bloom in 2014: linkages between biology and the water supply shutdown of Toledo, OH. *Environmental science & technology*, *51*(12), pp.6745-6755.

.

Trumpickas, J., Shuter, B.J. and Minns, C.K., 2009. Forecasting impacts of climate change on Great Lakes surface water temperatures. Journal of Great Lakes Research, 35(3), pp.454-463.

U.S. Geological Survey, 2017, National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), accessed February 6, 2017, at URL http://waterdata.usgs.gov/nwis/.

Vaccaro, L., Read, J. and Scavia, D., 2009. Michigan's Great Lakes Jobs. Ann Arbor, MI: Michigan Sea Grant. Available at: http://www. miseagrant. umich. edu/downloads/coastal/economy/09-101-Jobs-Report. pdf, 4.

Valipour, R., Boegman, L., Bouffard, D. and Rao, Y.R., 2017. Sediment resuspension mechanisms and their contributions to high-turbidity events in a large lake. Limnology and Oceanography, 62(3), pp.1045-1065.

Water Survey of Canada. 2017. HYDAT Database. Environment Canada: Ottawa, Canada. Extracted from https://www.canada.ca/en/environment-climate-change/services/water-overview/quantity/monitoring/survey/data-products-services/national-archive-hydat.html in September, 2017.

Woolway, R.I. and Merchant, C.J., 2018. Intralake heterogeneity of thermal responses to climate change: A study of large Northern Hemisphere lakes. Journal of Geophysical Research: Atmospheres, 123(6), pp.3087-3098.

Wuebbles, D., Cardinale, B., Cherkauer, K., Davidson-Arnott, R., Hellmann, J.J., Infante, D., Johnson, L., de Loe, R., Lofgren, B., Packman, A. and Selenieks, F., 2019. An Assessment of the Impacts of Climate Change on the Great Lakes. *The Environmental Law & Policy Center: Chicago, IL, USA*. http://elpc.org/wp-content/uploads/2019/03/Great-Lakes-Climate-Change-Report.pdf

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C. and Dyble, J., 2010. Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and meteorological data. Limnology and Oceanography, 55(5), pp.2025-2036.

Yarnal, B., 1993. Synoptic climatology in environmental analysis: a primer. Belhaven.

Yousef, F., Shuchman, R., Sayers, M., Fahnenstiel, G. and Henareh, A., 2017. Water clarity of the Upper Great Lakes: Tracking changes between 1998–2012. Journal of Great Lakes Research, 43(2), pp.239-247.

TABLES

Table 1 – Model performance by region. Open (rho) spearman correlations between NARX models (on open-loop) and observed KDI, Closed (rho) is the same, except for closed-loop NARX models. CL HR (80%) and CL HR (20%) are the hit rates for extreme KDI events (defined as days on which the closed-loop modeled KDI and observed KDI were at either the 80th or 20th percentiles of their respective time series). R2%Imp is the percentage-point improvement of the R-squared statistic between the final NARX models and the observations versus that of either a seasonal-trend model (R2%Imp Cycle) or a persistence model (R2%Imp AutoC) – the former is calculated using the closed-loop NARX models, the latter is calculated using open-loop models. Delta KDI is the change in KDI from 1979-2015 using a linear regression.

		OPEN (rho)	CLOSED (rho)	CL HR (80%)	CL HR (20%)	R2%Imp Cycle	R2%Imp AutoC	delta KDI
Superior	**R1**	0.67	0.20	25%	29%	0%	22%	-0.4
	R2	0.63	0.17	29%	21%	2%	21%	0.0
	R3	0.67	0.41	39%	42%	14%	16%	0.1
	R4	0.61	0.28	40%	22%	7%	17%	-0.1
	R5	0.59	0.21	37%	26%	3%	18%	-0.3
Huron	**R6**	0.88	0.64	50%	44%	6%	11%	-1.4
	R7	0.74	0.15	21%	24%	0%	48%	0.1
	R8	0.86	0.70	59%	49%	22%	13%	-1.7
	R9	0.91	0.76	73%	53%	12%	11%	-2.7
	R10	0.80	0.50	42%	42%	4%	20%	-1.2
	R11	0.84	0.37	35%	33%	6%	30%	-0.8
Erie	**R12**	0.86	0.34	32%	34%	4%	27%	-0.7
	R13	0.82	0.32	27%	32%	5%	18%	-0.1
	R14	0.79	0.31	23%	34%	5%	26%	0.6
Ont.	**R15**	0.69	0.20	30%	25%	0%	22%	-0.3
Michigan	**R16**	0.67	0.28	27%	23%	0%	26%	-1.0
	R17	0.84	0.68	69%	50%	5%	14%	-3.1
	R18	0.85	0.65	69%	49%	6%	12%	-2.6
	R19	0.89	0.61	54%	46%	3%	13%	-2.6
	R20	0.92	0.79	76%	60%	5%	8%	-2.9

Table 1 grayscale for print version

Table 1 – Model performance by region. Open (rho) spearman correlations between NARX models (on open-loop) and observed KDI, Closed (rho) is the same, except for closed-loop NARX models. CL HR (80%) and CL HR (20%) are the hit rates for extreme KDI events (defined as days on which the closed-loop modeled KDI and observed KDI were at either the 80th or 20th percentiles of their respective time series). R2%Imp is the percentage-point improvement of the R-squared statistic between the final NARX models and the observations versus that of either a seasonal-trend model (R2%Imp Cycle) or a persistence model (R2%Imp AutoC) – the former is calculated using the closed-loop NARX models, the latter is calculated using open-loop models. Delta KDI is the change in KDI from 1979-2015 using a linear regression.

		OPEN (rho)	CLOSED (rho)	CL HR (80%)	CL HR (20%)	R2%Imp Cycle	R2%Imp AutoC	delta KDI
Superior	**R1**	0.67	0.20	25%	29%	0%	22%	-0.4
	R2	0.63	0.17	29%	21%	2%	21%	0.0
	R3	0.67	0.41	39%	42%	14%	16%	0.1
	R4	0.61	0.28	40%	22%	7%	17%	-0.1
	R5	0.59	0.21	37%	26%	3%	18%	-0.3
Huron	**R6**	0.88	0.64	50%	44%	6%	11%	-1.4
	R7	0.74	0.15	21%	24%	0%	48%	0.1
	R8	0.86	0.70	59%	49%	22%	13%	-1.7
	R9	0.91	0.76	73%	53%	12%	11%	-2.7
	R10	0.80	0.50	42%	42%	4%	20%	-1.2
	R11	0.84	0.37	35%	33%	6%	30%	-0.8
Erie	**R12**	0.86	0.34	32%	34%	4%	27%	-0.7
	R13	0.82	0.32	27%	32%	5%	18%	-0.1
	R14	0.79	0.31	23%	34%	5%	26%	0.6
Ont.	**R15**	0.69	0.20	30%	25%	0%	22%	-0.3
Michigan	**R16**	0.67	0.28	27%	23%	0%	26%	-1.0
	R17	0.84	0.68	69%	50%	5%	14%	-3.1
	R18	0.85	0.65	69%	49%	6%	12%	-2.6
	R19	0.89	0.61	54%	46%	3%	13%	-2.6
	R20	0.92	0.79	76%	60%	5%	8%	-2.9

Table 2 – Averaged (across regions) monthly performance of NARX models on OL (open) and CL (closed) using spearman correlations.

	OPEN	CLOSED
JAN	0.81	0.33
FEB	0.77	0.28
MAR	0.76	0.41
APR	0.79	0.48
MAY	0.80	0.50
JUN	0.77	0.43
JUL	0.77	0.40
AUG	0.76	0.40
SEP	0.75	0.33
OCT	0.65	0.26
NOV	0.76	0.30
DEC	0.79	0.34

Table 2 grayscale for print version

Table 2 – Averaged (across regions) monthly performance of NARX models on OL (open) and CL (closed) using spearman correlations.

	OPEN	CLOSED
JAN	0.81	0.33
FEB	0.77	0.28
MAR	0.76	0.41
APR	0.79	0.48
MAY	0.80	0.50
JUN	0.77	0.43
JUL	0.77	0.40
AUG	0.76	0.40
SEP	0.75	0.33
OCT	0.65	0.26
NOV	0.76	0.30
DEC	0.79	0.34

Table 3 – Range of interannual variability in NARX model performance for each region, 1998-2015. Open is for OL models, closed is for CL models, min is the minimum annual spearman correlation for any calendar year, max is the maximum annual Spearman correlation for any calendar year.

REG	**OPEN**		**CLOSED**	
	MIN	**MAX**	**MIN**	**MAX**
1	0.41	0.74	-0.40	0.57
2	0.21	0.75	-0.27	0.57
3	0.20	0.80	-0.27	0.58
4	0.20	0.75	-0.05	0.57
5	0.20	0.74	-0.06	0.49
6	0.56	0.91	-0.34	0.72
7	0.48	0.95	-0.23	0.62
8	0.45	0.89	-0.12	0.77
9	0.49	0.89	-0.22	0.82
10	0.56	0.82	0.10	0.59
11	0.67	0.91	-0.32	0.56
12	0.70	0.89	-0.06	0.53
13	0.62	0.88	-0.15	0.58
14	0.66	0.86	-0.13	0.42
15	0.38	0.83	-0.08	0.43
16	0.44	0.71	-0.18	0.39
17	0.41	0.83	0.08	0.60
18	0.35	0.91	-0.21	0.68
19	0.60	0.91	-0.07	0.62
20	0.59	0.92	0.01	0.70

Table 4 – Region-by-region relative importance of the climate (and discharge) variables in the NARX model ensembles. All percentages are relative to each other (i.e. each row adds up to 100%). Each column is representative of the *average* importance of each of its components (e.g. GWTC has 11 weather types, so the GWTC column is the average importance of each weather type).

VAR	GWTC	500z	850T	MSLP	WIND	500pc	850pc	SLPpc	WNDpc	DIS
R1	8.4%	8.9%	9.1%	7.9%	8.4%	12.3%	11.6%	14.0%	14.1%	5.5%
R2	7.1%	8.7%	8.4%	7.3%	7.5%	12.2%	10.6%	12.4%	12.2%	13.6%
R3	8.0%	8.9%	8.9%	7.6%	7.8%	13.3%	12.1%	13.3%	13.2%	6.8%
R4	8.3%	9.0%	9.1%	7.8%	8.2%	12.9%	11.5%	12.5%	12.4%	8.2%
R5	7.1%	7.9%	8.3%	6.8%	7.1%	12.8%	11.2%	12.5%	12.1%	14.2%
R6	7.4%	9.0%	9.4%	6.6%	6.5%	13.0%	10.3%	10.7%	9.9%	17.1%
R7	7.6%	8.4%	9.2%	7.0%	7.1%	12.1%	11.7%	11.2%	11.3%	14.4%
R8	7.6%	9.0%	9.5%	7.0%	7.1%	12.6%	10.5%	11.4%	10.5%	15.0%
R9	6.6%	7.3%	8.1%	5.4%	5.5%	14.2%	11.0%	11.0%	9.5%	21.4%
R10	6.9%	7.5%	8.3%	6.2%	6.8%	13.5%	11.5%	13.2%	12.5%	13.8%
R11	7.4%	8.0%	8.4%	6.7%	6.9%	13.3%	11.9%	12.5%	12.2%	12.7%
R12	8.4%	8.7%	8.9%	7.4%	7.7%	12.9%	11.9%	12.8%	12.6%	8.9%
R13	8.5%	8.9%	9.3%	7.8%	8.4%	13.0%	11.5%	13.0%	13.1%	6.5%
R14	8.0%	8.8%	9.1%	7.2%	7.5%	13.0%	10.9%	12.5%	11.8%	11.3%
R15	8.7%	8.8%	8.4%	8.0%	8.4%	12.8%	11.8%	13.4%	13.6%	6.0%
R16	8.0%	8.0%	7.6%	6.9%	7.5%	14.7%	12.1%	13.5%	12.2%	9.6%
R17	8.0%	8.2%	8.1%	6.7%	7.1%	14.0%	10.9%	13.1%	11.9%	11.9%
R18	7.8%	8.6%	8.1%	6.7%	7.0%	13.8%	9.8%	12.5%	11.4%	14.4%
R19	8.8%	8.1%	8.2%	7.2%	7.8%	12.4%	10.9%	13.4%	13.5%	9.5%
R20	7.0%	6.9%	6.9%	5.2%	5.4%	14.0%	10.0%	11.6%	9.7%	23.2%
AVG	7.8%	8.4%	8.6%	7.0%	7.3%	13.1%	11.2%	12.5%	12.0%	12.2%

Table 4 grayscale for print version

Table 4 – Region-by-region relative importance of the climate (and discharge) variables in the NARX model ensembles. All percentages are relative to each other (i.e. each row adds up to 100%). Each column is representative of the *average* importance of each of its components (e.g. GWTC has 11 weather types, so the GWTC column is the average importance of each weather type).

VAR	GWTC	500z	850T	MSLP	WIND	500pc	850pc	SLPpc	WNDpc	DIS
R1	8.4%	8.9%	9.1%	7.9%	8.4%	12.3%	11.6%	14.0%	14.1%	5.5%
R2	7.1%	8.7%	8.4%	7.3%	7.5%	12.2%	10.6%	12.4%	12.2%	13.6%
R3	8.0%	8.9%	8.9%	7.6%	7.8%	13.3%	12.1%	13.3%	13.2%	6.8%
R4	8.3%	9.0%	9.1%	7.8%	8.2%	12.9%	11.5%	12.5%	12.4%	8.2%
R5	7.1%	7.9%	8.3%	6.8%	7.1%	12.8%	11.2%	12.5%	12.1%	14.2%
R6	7.4%	9.0%	9.4%	6.6%	6.5%	13.0%	10.3%	10.7%	9.9%	17.1%
R7	7.6%	8.4%	9.2%	7.0%	7.1%	12.1%	11.7%	11.2%	11.3%	14.4%
R8	7.6%	9.0%	9.5%	7.0%	7.1%	12.6%	10.5%	11.4%	10.5%	15.0%
R9	6.6%	7.3%	8.1%	5.4%	5.5%	14.2%	11.0%	11.0%	9.5%	21.4%
R10	6.9%	7.5%	8.3%	6.2%	6.8%	13.5%	11.5%	13.2%	12.5%	13.8%
R11	7.4%	8.0%	8.4%	6.7%	6.9%	13.3%	11.9%	12.5%	12.2%	12.7%
R12	8.4%	8.7%	8.9%	7.4%	7.7%	12.9%	11.9%	12.8%	12.6%	8.9%
R13	8.5%	8.9%	9.3%	7.8%	8.4%	13.0%	11.5%	13.0%	13.1%	6.5%
R14	8.0%	8.8%	9.1%	7.2%	7.5%	13.0%	10.9%	12.5%	11.8%	11.3%
R15	8.7%	8.8%	8.4%	8.0%	8.4%	12.8%	11.8%	13.4%	13.6%	6.0%
R16	8.0%	8.0%	7.6%	6.9%	7.5%	14.7%	12.1%	13.5%	12.2%	9.6%
R17	8.0%	8.2%	8.1%	6.7%	7.1%	14.0%	10.9%	13.1%	11.9%	11.9%
R18	7.8%	8.6%	8.1%	6.7%	7.0%	13.8%	9.8%	12.5%	11.4%	14.4%
R19	8.8%	8.1%	8.2%	7.2%	7.8%	12.4%	10.9%	13.4%	13.5%	9.5%
R20	7.0%	6.9%	6.9%	5.2%	5.4%	14.0%	10.0%	11.6%	9.7%	23.2%
AVG	7.8%	8.4%	8.6%	7.0%	7.3%	13.1%	11.2%	12.5%	12.0%	12.2%

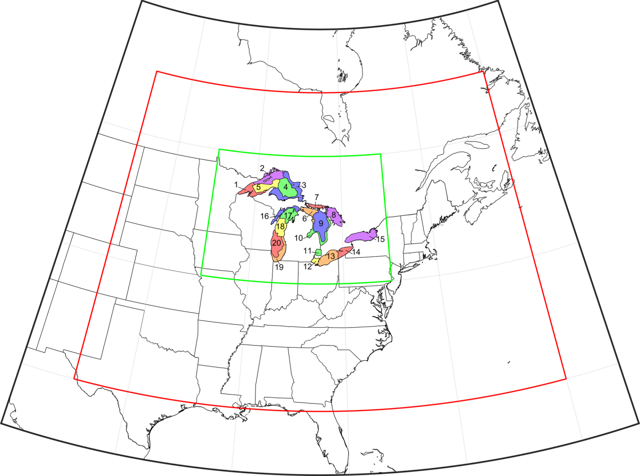
FIGURE CAPTIONS

Figure 1 – The 20 regions within which KDI is defined, and the spatial domain used for the classification of CPs (boxes). The larger (red box) domain is used for 500Z and SLP; the smaller (green box) is used for 10-m winds and 850T.

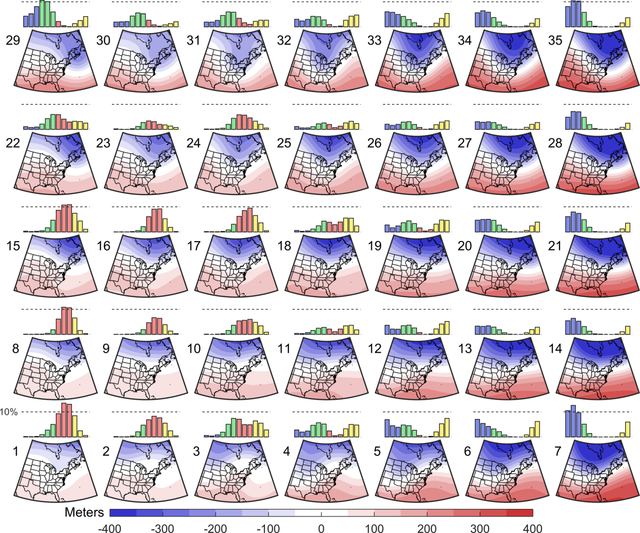
Figure 2 – The SOM of CPs created for anomalous 500Z data (units are in meters). Above each map is a bar graph showing monthly frequency of occurrence (from December on the left to November on the right), with blue bars being winter months (December, January, February), green bars being spring months (March-May), orange bars being summer months (June-August), and yellow bars being autumn (September-November). The dotted line above each indicates a 10% frequency of occurrence within a month. Figure modified from: Smith, E.T., Lee, C.C., Barnes, B.B., Adams, R.E., Pirhalla, D.E., Ransibrahmanakul, V., Hu, C., Sheridan, S.C., 2020. A synoptic climatological analysis of the atmospheric drivers of water clarity variability in the Great Lakes. DOI: 10.1175/JAMC-D-19-0156.1. © American Meteorological Society. Used with permission. Accepted for publication in the Journal of Applied Climatology and Meteorology.

Figure 3 – Region-by-region monthly time series of NARX models (blue) and observed KDI (orange), 1979-2015 (x-axis). Note that the y-axis (KDI) varies in each subplot/region.

Figure 4 – Region-by-region reforecast performance (blue line) vs. *x*-day autocorrelation (orange line), 2016-2018. Performance is calculated using Pearson correlations (y-axis) between NARX modeled KDI at *x*=1 to *x*=60 days (*x*-axis of each figure) of lead time and observed KDI.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20



29
30
31
32
33
34
35
22
23
24
25
26
27
28
15
16
17
18
19
20
21
8
9
10
11
12
13
14
10%
1
2
3
4
5
6
7
Meters
-400
-300
-200
-100
0
100
200
300
400