Response of Benthic Macroinvertebrate Communities to Climate Change

Abstract

Aquatic ecosystems are impacted by climate change. However, the level of impacts, especially on macroinvertebrates, are not well known, which is the goal of this study. Here, three biotic indices were used. The Saginaw River Basin, the largest watershed in Michigan was selected for this study. Since no known modeling package exists to evaluate the impacts of climate change on these biotic indices, a coupled modeling approach was proposed. First, the Soil and Water Assessment Tool (SWAT) was setup to estimate streamflow and water quality for all stream segments. Next, the results from the SWAT model in combination with long-term biological monitoring data were used to develop biotic indices predictive models using a two-phase approach. An assemblage of 42 climate scenarios (2040-2060) were obtained from the National Aeronautics and Space Administration, which are biased corrected and statistically downscaled. The dataset provides two extreme Representative Concentration Pathways (RCPs) to account for best (RCP 4.5) and worst (RCP 8.5) case scenarios. The climate data were incorporated in SWAT and biotic indices predictive models to understand stream health conditions in the watershed. Results showed that the majority of streams have a probability of declining in values of biotic indices while streams that currently have excellent conditions are the most vulnerable streams to the climate change. This modeling approach provides a new mean to link more strongly biota indicators with flow regime indicators, which is of prime interest for the implementation of ecohydrology principles.

Keywords: Climate change; Stream health; Macroinvertebrates; Saginaw; SWAT

1. Introduction

Climate change is one of the largest threats to modern society, impacting all aspects of civilization from agricultural production to freshwater availability due to increase in climate variability (Grimm et al., 2013). As temperatures continue to rise and the number of extreme hydrological events (i.e. droughts and floods) increases, the surrounding environment in which species are living become inhospitable due to increased competition for limited resources, which can even result in species extinction (Zhu et al., 2013). In fact, anthropogenic activities and climate change may cause the world's sixth mass extinction (Barnosky et al., 2011). In order to mitigate the negative impacts of global climate change, it has become necessary to understand how climate change will affect different aspects of ecosystems (Knight and Harrison, 2013).

Considering freshwater ecosystems, there have been concerns that changes in precipitation patterns and temperature rise will have major impacts on species living in this environment (Isaak et al., 2012; Arnell and Gosling, 2013; Beniston and Stoffel, 2014). For example, changes in precipitation can alter flow patterns in rivers and streams (Nejadhashemi et al., 2012; Arnell and Gosling, 2013; Beniston and Stoffel, 2014), which in turn impacts organism communities resulting in loss of biodiversity. In addition, rising water temperatures alter river and stream

thermal regimes, which impacts organism reproduction and lifecycles, especially for cold-water species like salmonids (Isaak et al., 2012; Beniston and Stoffel, 2014). Despite the importance of findings from these studies among others (Walther et al., 2002; Harley et al., 2006; Isaak et al., 2012; Lurgi et al., 2012; Grimm et al., 2013), they mainly relied on field sampling that can only capture past climate impacts on freshwater ecosystems. One approach that can overcome these limitations is the use of models (Giri, et al., 2012) to predict biotic indices beyond the sampling points (Einheuser et al., 2012; Woznicki et al., 2016a). Then the biotic index model can be integrated with the climate models, particularly with Global Climate Models (GCMs), to simulate future climate scenarios. GCMs are a collection of physically based numerical models that simulate the impacts of greenhouse gas emissions on the global climate system (IPCC, 2013). However due to the coarse resolution of GCMs, downscaling methods are needed to create local climate projections, which can be used for climate change impact studies (Schmatz et al., 2015). One of the most commonly used downscaling methods is the delta change factor method (Schmatz et al., 2015; Woznicki et al., 2016b). The delta change factor method is based on three assumptions, which include: (1) GCMs are better suited for simulating relative changes rather than absolute changes; (2) GCM biases are similar across future and control periods; and (3) the variability in climate patterns is constant between current and future time periods (Woznicki et al., 2016b). However, this method is limited by its assumptions, especially the assumption that climate variability will remain constant. Therefore, in this study, GCM data was statistically downscaled and bias-corrected to better represent future climate conditions.

The main goal of this study is to evaluate the impacts of climate change on stream health based on three benthic macroinvertebrate indices. The specific objectives of this study were to: 1) develop stream health predictive models for biotic indices and 2) evaluate the future impacts of climate change on benthic macroinvertebrate communities using the concept of probability.

2. Materials and Methods

2.1 Study area

The area of study is called the Saginaw River Basin, which is the largest watershed in Michigan with a six-digit Hydrological Unit Code number of 040802 (Figure 1). The total watershed drainage area is 16,120 km² that drains to Lake Huron. The US Environmental Protection Agency (USEPA) identified the Saginaw River Basin as an area of concern in the Great Lakes Basin due to sediment contamination, fisheries degradation, and loss of recreational values (EPA, 2015). Agricultural and Forest cover are dominant land uses (74%), while water/wetlands and urban area cover 14% and 12% of the watershed area, respectively.

Figure 1.

2.2 Biological Data

In order to insure that freshwater ecosystems can be used by future generations it has become necessary to monitor stream conditions to guarantee their sustainability. One of the most common techniques for monitoring stream conditions is the use of bioassessments, which take

into account a stream's chemical, physical, and biological components to determine the overall condition within the stream (Brazner et al., 2007; Pelletier et al., 2012; Leigh et al., 2013). Bioassessments can be based on a variety of organisms including algae, amphibians, diatoms, fish, macroinvertebrates, mammals, microorganisms, periphyton, phytoplankton, plants, reptiles, and zooplankton (Herman and Nejadhashemi, 2015). In this study, three macroinvertebrate indices, including: the Hilsenhoff Biotic Index (HBI), the Family Index of Biological Integrity (FIBI), and the EPT index which represents the degradation of *Ephemeroptera* (mayflies), *Plecoptera* (stoneflies), and *Trichoptera* (caddisflies) (Lenat, 1988) were used. The HBI index varies from 0 (the best condition) to 10 (the worst condition). The HBI represents the response of aquatic species to organic pollutants (Hilsenhoff, 1988). The FIBI ranges from 0 (the worst condition) to 45 (the best condition), and represents the composition of species in family taxonomic resolution with respect to pollution tolerance (Kerans and Karr, 1994; Herman and Nejadhashemi, 2015). An EPT score of zero represents the worst stream health condition while higher values represent better stream health conditions but the upper limits are different in different regions (Woznicki et al., 2015; Herman and Nejadhashemi, 2015). These scores were reported based on sampling performed by the Michigan Department of Environmental Quality (MDEQ) between June 1996 to September 2003 (MDEQ, 1997). Figure 2 shows the location of 308 monitoring sites in the study area.

Figure 2.

2.3 Modeling Process

In order to evaluate the impacts caused by climate change on macroinvertebrate communities beyond the biological sampling points, stream health predictive models should be developed. Figure 3 illustrates the modeling process. The first step in the model development is to assemble physiographical (e.g. topography, land use, soil) and climatological (e.g. precipitation and temperature) data in order to setup a watershed model. The Soil and Water Assessment Tool (SWAT) is the watershed model selected for this purpose. The calibrated SWAT model simulates water quality and quantity parameters for all streams on a daily basis. The stream flow estimations from SWAT were used to calculate 171 biologically related hydrological indices using the MATLAB Hydrological Index Tool (MHIT) model. The sediment and nutrients loads estimated from SWAT model along with biologically related hydrological indices calculated by MHIT were then used to develop three predictive stream health models (HBI, FIBI, EPT) using the two-phase approach introduced by Abouali et al. (2016a). In the first step, both of the biotic index and associated error are estimated while in the second step, these initial estimates are used in development of the final predictive model. Using the developed predictive models, the biologicals measures of 13,831 stream segments within the study area were calculated for current/historical (1980-2000) and 42 future (2040-2060) climate scenarios (Table 1).

Figure 3.

2.3.1 Physiographical/Climatological Datasets

The first step in the modeling process is to obtain physiographical and climatological datasets for setting-up the watershed model (SWAT). Concerning the physiographical datasets, a topography map, including digital elevation in 1/3 arc-second spatial resolution was downloaded from the U.S. Geological Survey (USGS) national elevation dataset (NED, 2015). The Cropland Data Layer (CDL) of Michigan was obtained from the U.S. Department of Agriculture (USDA)-National Agricultural Statistics Service (NASS) (NASS, 2014). Soil Survey Geographic Database (SSURGO) was acquired from the USDA-National Resources Conservation Service (NRCS) that includes physical and chemical soil properties (SSURGO, 2014).

Regarding the observed climatological data, daily precipitation and minimum and maximum temperatures were obtained from the National Climate Data Center (NCDC) for 16 stations (Figure 2) within the study area (NCDC, 2015). In addition, historical and future modeled climatological datasets were obtained from the National Aeronautics and Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP, 2015). This dataset is bias-corrected and statistically downscaled using the Bias-Correction Spatial Disaggregation (BCSD) method (Thrasher et al., 2013). NEX-GDDP is comprised of 21 General Circulation Models (GCM) conducted under the Coupled Model Intercomparison Project Phase 5 (CMIP5) (Thrasher et al., 2013). These datasets are available for two Representative Concentration Pathways (RCPs) including RCP 4.5 and RCP 8.5 at the resolution of 0.25 degrees (about 25 km × 25 km) (Thrasher et al., 2013). The modeled historical (1980 to 2000) and future (2040 to 2060) precipitation as well as minimum and maximum temperatures were obtained from the NEX-GDDP. The offset of historical climate trends and observed values obtained from weather stations were adjusted by applying additional downscaling (Woznicki et al., 2014). Table S1 shows the mean and variance of future climate conditions from each of the 21 climate models for the locations of 16 weather stations while Table 1 shows the range. The range of precipitation and maximum and minimum temperatures for current condition and based on observed data from the same 16 weather stations are (778 - 843 mm), (12.62-14.37 °C), and (0.38-3.60 °C), respectively.

Table 1.

2.3.2 Soil and Water Assessment Tool

The physiographical and climatological datasets obtained in previous section were used to setup a watershed model. The Soil and Water Assessment Tool (SWAT revision 2012) was used in this study to simulate streamflows, sediment, and nutrient loads for all stream segments within the study area. SWAT was developed by USDA-Agricultural Research Service (ARS). The model was design to evaluate the impacts of land management practices and climate variabilities on streamflow, sediment transport, and agrochemical yields in large and diverse watersheds (Gassman et al., 2007; Neitsch et al., 2011). SWAT uses the modified Soil Conservation Service curve number method for estimating surface runoff while the modified Universal Soil Loss Equation (MUSLE) is used in the model to estimate eroded soil and sediment delivery. In addition, SWAT considers three different forms for nitrogen (organic, mineral held by soil, and soluble mineral) and phosphorus (organic, insoluble mineral, and soluble in soil but plant-available) for modeling purposes (Neitsch et al., 2011).

The watershed in this study has 13,831 reaches and subwatersheds that were mapped as predefined layers to SWAT. The layers were developed by Michigan Institute of Fishery Research. The model was run for two time periods of current (1980-2000) and future (2040-2060) climate scenarios. However, the first three years of simulations for each period (1978-1980 and 2038-2040) were considered as warmup periods.

The SWAT model was calibrated and validated for stream flow, sediment, nitrogen and phosphorus loads according to the model evaluation criteria introduced by Moriasi et al. (2007). These criteria include: 1) Nash-Sutcliffe efficiency (NSE) representing residual variance ratio to the observed data variance; 2) percent bias (PBIAS) showing by how much the simulated results are larger or smaller than the observed values; and 3) ratio of the root mean square error to the standard deviation of observed data (RSR) (Moriasi et al., 2007). Stream flow data are available on daily basis from nine USGS streamflow stations (04144500, 04147500, 04148140, 04148500, 04151500, 04154000, 04155500, 04156000, 04157005) within the region (Figure 1) from 2001 to 2010. The streamflow datasets were used for SWAT model calibration and validation for ten years (five years each). Sediment, nitrogen and phosphorous loads were calibrated and validated for eight years (four years each) from 1998 to 2005 on a monthly basis for an EPA water quality monitoring site (90177), which is located close to the mouth of the watershed (EPA STORET, 2015) (Figure 1). The USGS Load Estimator (LOADEST) was used to create monthly loads based on daily observed data. LOADEST develops a regression model to estimate pollution loads based on streamflow time series and limited constituent concentration measurements (Runkel et al., 2004). Using three statistical estimation methods including the Adjusted Maximum Likelihood Estimation (AMLE), Maximum Likelihood Estimation (MLE), and Least Absolute Deviation (LAD), LOADEST develops three regression models. Two of them are recommended when model residuals are normally distributed, while AMLE can also be used for censored data sets. However, the third one (LAD) is good for not normal distributed residuals (Runkel et al., 2004).

2.4 Stream health models

MHIT developed by Abouali et al. (2016b) was used to calculate 171 biologically relevant hydrological indices. The MHIT model is a modified version of the hydrological index tool that was developed by USGS and uses long-term stream flow data that in this study was obtained from SWAT. MHIT indices represent the magnitude, frequency, duration, timing, and rate of changes of discharge conditions (Poff et al., 1997; Henriksen et al., 2006). In addition to average, minimum, and maximum, MHIT use the concept of exceedance and percentile to calculate the indices. The detailed descriptions and relevant equations are provided by Henriksen et al. (2006).

Then the two-phase approach developed by Abouali et al. (2016a) was implemented to develop predictive models for each of the biotic indices (FIBI, HBI, and EPT).

The two-phase approach can be summarized as follows (Abouali et al., 2016a):

- **Phase I:** Initial estimate of the dependent variable and the associated error.
 - **Part I:** A predictive model is developed to estimate initial index values ($\hat{Y}(P)$).

- **Part II:** A predictive model is developed to estimate errors of the initial estimates ($\hat{\varepsilon}(\hat{P})$).
- **Phase II:** Establish a predictive model using $\hat{Y}(P)$ and $\hat{\varepsilon}(\hat{P})$ as inputs to estimate the final index values (Y).

For this study and in the first phase, Partial Least Square Regression (PLSR) was used to estimate initial values for the index and the associated error. PLSR projects both the dependent and independent variables into a new space formed by mutually orthogonal components; hence, unlike regular linear regression, PLSR is not affected by the multicollinearity among the variables. This feature of PLSR also eliminates the need for input parameter selection; hence, PLSR is not limited in the number of input parameters. In fact, PLSR was designed for situations where there are more input parameters than there are measurements. In such a condition, regular linear regression fails. For the second phase, the Adaptive Network Fuzzy Inference System (ANFIS) was used to predict the index using the initial estimates from the first phase. ANFIS is a non-linear approach and combines Artificial Neural Network (ANN) and Fuzzy Logic framework. ANFIS has shown to be suitable for ecological analysis (Einheuser et al., 2013a; Einheuser et al., 2013b; Woznicki et al., 2015); however, it is limited in the number of input parameters that it can use, due to both computational resources and the available measurements that it needs (Hamaamin et al., 2013; Woznicki et al., 2015).

According to the River Continuum Concept (Vannote et al., 1980), the occurrence of organisms varies within the stream network. However, the network can be generally classified into headwater (orders 1-3), medium-sized (orders 4-6), and large (7-10). Since the maximum stream order in Saginaw River Basin was 7, in this study three sets of models were developed for each biotic index and grouped into ‘all’ that contains all streams in the watershed, “headwater” that includes all streams with orders 1 to 3, and ‘medium-sized and large’ that includes all streams with orders 4 to 7.

Abouali et al. (2016a) developed a MATLAB package implementing the two-phase predictive model approach. The package can be accessed free of charge for academic and non-commercial purposes at: https://github.com/mabouali/PLSR_ANFIS. ANFIS make use of fuzzy membership function for input variables, which varies from 0 (no membership) and 1 (full membership). The MATLAB package allows for Gaussian curve, Gaussian combination, Generalized bell-shaped, Triangular-shaped, and Trapezoidal-shaped membership functions. These membership functions are commonly used in ecological studies (Hamaamin et al., 2014; Woznicki et al., 2015; Herman et al., 2016; Woznicki 2016b). For the output, there are only two options for the membership functions: linear and constant. The MATLAB package searches among all the possible combinations automatically. In all phases, 10-fold cross validation is used and the model with the lowest Root Mean Square Error (RMSE) and the highest correlation is chosen among all the possible combinations. For each biotic index predictive model, the overall performance was evaluated against observed data obtained from macroinvertebrate sampling stations (Figure 2). Considering three biotic indices and three groups of stream sizes (all, headwater, and medium-sized and large), in the first phase, one PLSR model is developed for the initial estimate of the index, and another PLSR model is developed to estimate the associated error. In the second

phase, two input and two output parameters were considered for the ANFIS models. Five different membership functions are used for the input variables while two different membership functions are used for the output. Therefore, 90 ($3 \times 3 \times 5 \times 2$) different ANFIS models are trained and developed for each biotic index. Since 10-fold cross validation is used at all steps/phases; hence, everything is repeated 10 times.

2.5 Stream health models

In order to evaluate the impacts of climate change on benthic macroinvertebrate communities, the probability of decline was calculated using current and 42 climate scenarios. The probability of decline of each benthic macroinvertebrate was calculated by comparing the current and future simulations using the Weibull (1939) equation:

$$P = \frac{m}{n+1} \quad (1)$$

where, P is the probability of decline for each benthic macroinvertebrate being less or equal to current value, m is the ranked of the current value for the benthic macroinvertebrate, and n is the total number of observations. Calculated probability of declines were finally grouped into four groups (0-25%, >25–50%, >50–75%, and >75–100%).

3. Results and Discussion

3.1 SWAT model calibration/validation

Stream flow was calibrated and validated for ten years (2001-2010) on a daily basis. According to Moriasi et al. (2007), a model performance can be considered satisfactory if NSE is above 0.5, PBIAS ranges between +/-25% and RSR is below 0.7. A summary of calibration and validation criteria of the nine USGS streamflow stations (Figure 1) are provided in Table 2. According to this criteria and results from Table 2, the hydrologic components of the model were performing satisfactory.

Table 2.

The loads were used for water quality calibration and validation. For evaluating water quality performance of a watershed model, Moriasi et al. (2007) considered similar NSE and RSR values for streamflow, while satisfactory range of PBIAS for sediment and nutrients (nitrogen and phosphorous) were +/-55 and +/-70, respectively. Table 3 shows a summary of water quality calibration and validation results. According to the aforementioned criteria and results from Table 3, the SWAT model performance for water quality parameters is satisfactory.

Table 3.

3.2 Stream health models development

Table 4 presents the summary of the two-phase models developed for the HBI, FIBI and EPT. As can be seen in Table 4, the models developed for headwater streams (orders 1-3) and medium-sized and larger streams (orders 4-7) have better predictions (higher R^2 , lower RMSE) than the models developed for all of the streams. This is consistent with previous studies on stream health modeling (Woznicki et al., 2015) in which models developed using stream order or k -means

clustering outperformed the model developed for the entire watershed. Therefore, the stream models developed for stream orders 1 to 3 and 4 to 7 were selected to evaluate stream health conditions under current and future climate scenarios. In general, the predictabilities of stream health models were greatly improved compared with previous studies. For example, Herman et al. (2015) developed a stream health model using ANFIS technique for the Honeyoey Creek-Pine Creek Watershed in Michigan in which the model performances (R^2) for stream order 1-3 and >3 were 0.514 and 0.699, respectively. Einheuser et al. (2012) developed stream health models for the three benthic macroinvertebrate indices (HBI, FIBI, and EPT). Using stepwise linear regression technique, the R^2 -varies from 0.19 for HBI to 0.36 for EPT while for ANFIS R^2 -varies from 0.28 for FIBI to 0.39 for EPT.

Table 4.

Based on the biotic index score, the streams were categorized into five groups ranging from very poor to excellent (Figure S13 in the Supplementary Materials). Stream health classes for EPT, FIBI, and HBI were obtained from (Woznicki et al., 2015) in which EPT varies from 0 to >10, FIBI varies from 0 to 45, and HBI varies from 0 to 10. These ranges were observed within the study area from long-term monitoring. Based on sensitivity of benthic indices, the response varies in different part of the watershed (Figure S13). For example, the HBI, measuring the response of benthic macroinvertebrates communities to organic pollutants, showed that 28% of streams have excellent condition, 19% of streams fall in the fair and good class, and 48% were classified as very poor and poor conditions. For the FIBI, that measures the index of biotic integrity with respect to pollution tolerance, classified only 19% of streams as excellent, 27% of streams were classified as fair and good, and 49% were classified as very poor and poor conditions. Finally, for the EPT, measuring degradation of three species (mayflies, stoneflies, and caddisflies), 20% of streams were classified as excellent, 35% of streams were classified as fair and good, and 40% were classified as very poor and poor. However, it is important to note that the areas that identified as e.g. ‘excellent’ for one index are not necessarily categorized as ‘excellent’ for other indices.

In order to obtain the overall stream health condition, the values of the three macroinvertebrate indices were considered for each stream segment. However, since each of these indices has different scaling system and range, it is not possible to obtain the overall score by simply averaging the three macroinvertebrate measures. Under this circumstance, one should normalize the values of macroinvertebrate indices using the linear scaling technique (Esfahanian et al., 2017) with values varying from 0 (the worst condition) to 1 (the best condition) with 0.2 intervals for each class. Therefore, in the new system, scores varying from 0-0.2 fall in very poor class and so on. Then the three measure scores (HBI, FIBI, EPT) were averaged for each stream segment. Using this approach, an overall macroinvertebrate score for the current condition was developed that can be used instead of three indices with different ranges. Among 13,831 streams in the region, fair and good classes are dominant (54.1%), followed by poor and very poor classes (33.1%). Streams with excellent condition are rare (8.4%) and mostly limited to small upstream reaches.

Comparison of current stream health scores and land use map (Figure 4) shows that forest and water lands (wetlands and forested wetlands) have better stream health conditions than streams located within agricultural lands. More than 64% of streams with excellent condition and 52% of streams with good or fair condition are in forest and water lands. These streams are less affected by anthropogenic activities. While the majority of streams with poor or very poor condition (63%) are located in agricultural lands showing how human activities have negatively affected stream health. Besides that, moving from upstream to the outlet, the stream health condition gets worse and poor and very poor rivers become dominant (Arthington and Pearson, 2007).

Figure 4.

3.3 Future climate scenarios

Using 21 climate models with two RCPs (4.5 and 8.5) listed in Table 1, 42 climate scenarios were used to get three macroinvertebrates scores and overall scores as mentioned in section 3.2 for each stream segment. Compared to the current condition, it was found that many scenarios predicted worse conditions for each stream. Figure 5 shows the probability of decline for each stream segment or the percentage of future scenarios that predicted worse conditions for the given reach.

Figure 5.

As shown in Figure 5, only about one third of the streams have a low probability of decline (0%-25%). Compared to the current condition, almost all of streams with excellent conditions currently, have the highest probability of decline (75%-100%), meaning that there may be no reaches with excellent conditions in near future (2040-2060). Many of these reaches are small streams with low flows and are more vulnerable to climate change. Streams with very poor, poor or fair conditions have less probability of decline, since these streams have been already degraded. This study showed how vulnerable macroinvertebrate communities are to the climate change and almost all streams in this region will have fair or below (poor and very poor) condition for macroinvertebrates in the near future.

4. Conclusion

In general, the two-phase approach has significantly improved the performances of stream health models. For the same study area, the reported R^2 for best performance models (HBI, FIBI and EPT) were below 0.6 (Einheuser et al., 2012). However, in this study, using the two-phase approach improved R^2 to above 0.90.

Current overall stream health scores showed that most of the rivers (54%) have fair or good conditions followed by poor and very poor (33%) conditions for macroinvertebrates communities. Streams located in forestlands had excellent conditions, while rivers with very poor and poor conditions were mostly located in agricultural and urban lands. Regarding the future climate scenarios, the probability of decline map shows that many of the streams have a

probability of decline above 50%. Among them, forest streams that had excellent conditions in the current climate scenario were found to be the most vulnerable to climate change, ending up with almost no streams with excellent conditions for macroinvertebrates communities in near future (2040-2060). Meanwhile, streams with very poor, poor or fair conditions have less probability of decline. However, this does not mean that these streams are less vulnerable to climate change. The reason for lower probability of decline can be explained due to the fact that these streams have been already degraded; therefore, additional degradation does not translate to significant drop in stream health condition.

In order to better understand the impacts of climate change on stream health, future studies should also consider additional aquatic organisms such as fish. In addition, it is expected that climate change will impact terrestrial ecosystems that indirectly affect aquatic ecosystems. This study did not account for these impacts and possible human responses that are important to be included in future studies. Finally, this study did not consider the impacts of water temperature alteration due to climate change; therefore, it is recommended that temperature should be also incorporated for development of biotic indices predictive models.

5. Acknowledgment

This work is supported by the USDA National Institute of Food and Agriculture, Hatch project MICL02212. Climate scenarios used were from the NEX-DCP30 dataset, prepared by the Climate Analytics Group and NASA Ames Research Center using the NASA Earth Exchange, and distributed by the NASA Center for Climate Simulation (NCCS).

6. References

- Abouali M., Nejadhashemi, A. P., Daneshvar F., Woznicki, S. A. 2016a, Two-Phase approach to improve stream health modeling. *Ecol. Inform.* 34, 13-21.
- Abouali, M., Daneshvar, F., Nejadhashemi, A.P., 2016b. MATLAB Hydrological Index Tool (MHIT): A high performance library to calculate 171 ecologically relevant hydrological indices. *Ecol. Inform.* 33, 17-23.
- Arnell, N.W., Gosling, S.N., 2013. The impacts of climate change on river flow regimes at the global scale. *J. Hydrol.* 486, 351–364.
- Arthington, A.H., Pearson, R.G., 2007. Biological indicators of ecosystem health in wet tropics streams. Final Report to the Catchment to Reef Research Program. CRC Reef and Rainforest CRC. James Cook University.
- Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., Ferrer, E.A., 2011. Has the Earth's sixth mass extinction already arrived? *Nature* 471, 51–57.
- Beniston, M., Stoffel, M., 2014. Assessing the impacts of climatic change on mountain water resources. *Sci. Total Environ.* 493, 1129–1137.
- Brazner, J.C., Danz, N.P., Niemi, G.J., Regal, R.R., Trebitz, A.S., Howe, R.W., Hanowski, J.M., Johnson, L.B., Ciborowski, J.J.H., Johnston, C.A., Reavie, E.D., Brady, V.J., Sgro, G.V., 2007.

- Evaluation of geographic, geomorphic and human influences on Great Lakes wetland indicators: A multi-assemblage approach. *Ecol. Indic.* 7, 610–635.
- Einheuser, M.D., Nejadhashemi, A.P., Sowa, S.P., Wang, L., Hamaamin, Y.A., Woznicki, S.A., 2012. Modeling the effects of conservation practices on stream health. *Sci. Total Environ.* 435–436, 380–391.
- Einheuser, M.D., Nejadhashemi, A.P., Wang, L., Sowa, S.P., Woznicki, S.A., 2013a. Linking biological integrity and watershed models to assess the impacts of historical land use and climate changes on stream health. *Environ. Manage.* 51, 1147–1163.
- Einheuser, M.D., Nejadhashemi, A.P., Woznicki, S.A., 2013b. Simulating stream health sensitivity to landscape changes due to bioenergy crops expansion. *Biomass Bioenerg.* 58, 198–209.
- EPA, 2015. Saginaw River and Bay Area of Concern. URL: <http://www.epa.gov/saginaw-river-bay-aoc> (Last Accessed 2.14.16).
- EPA STORET, 2015, Storage and retrieval and water quality exchange. URL: <https://www.epa.gov/waterdata/storage-and-retrieval-and-water-quality-exchange> (Last Accessed 3.22.16).
- Esfahanian, E., Nejadhashemi, A.P., Abouali, M., Adhikari, U., Zhang, Z., Daneshvar, F., Herman, M. R. 2017. Development and evaluation of a comprehensive drought index. *J. Environ. Manage.* 185, 31–43.
- Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. *T. ASABE* 50, 1211–1250.
- Giri, S., Nejadhashemi, A.P., Woznicki, S.A., 2012. Evaluation of targeting methods for implementation of best management practices in the Saginaw River Watershed. *J. Environ. Manage.* 103, 24–40. 3
- Grimm, N.B., Chapin, F.S., Bierwagen, B., Gonzalez, P., Groffman, P.M., Luo, Y., Melton, F., Nadelhoffer, K., Pairis, A., Raymond, P.A., Schimel, J., Williamson, C.E., 2013. The impacts of climate change on ecosystem structure and function. *Front. Ecol. Environ.* 11, 474–482.
- Hamaamin, Y.A., Adhikari, U., Nejadhashemi, A.P., Harrigan, T., Reinhold, D.M., 2014. Modeling *Escherichia coli* removal in constructed wetlands under pulse loading. *Water Res.* 50, 441–454.
- Hamaamin, Y.A., Nejadhashemi, A.P., Einheuser, M.D., 2013. Application of fuzzy logic techniques in estimating the regional index flow for Michigan. *T. ASABE* 56, 103–115.
- Harley, C.D.G., Randall Hughes, A., Hultgren, K.M., Miner, B.G., Sorte, C.J.B., Thornber, C.S., Rodriguez, L.F., Tomanek, L., Williams, S.L., 2006. The impacts of climate change in coastal marine systems. *Ecol. Lett.* 9, 228–241.
- Henriksen, J.A., Heasley, J., Kennen, J.G., Nieswand, S., 2006. Users' manual for the hydroecological integrity assessment process software (including the New Jersey Assessment Tools). U. S. Geological Survey. Biological Resource Discipline, Open File Report 2006-1093, 71 p.
- Herman, M.R., Nejadhashemi, A.P., 2015. A review of macroinvertebrate- and fish-based stream health indices. *Ecohydrol. Hydrobiol.* 15, 53–67.

- Herman M.R., Nejadhashemi, A.P., Daneshvar F., Abouali M., Ross D.M., Woznicki, S. A., Zhang Z. 2016. Optimization of bioenergy crop selection and placement based on a stream health indicator using an evolutionary algorithm. *J. Environ. Manage.* 181, 413-424.
- Herman, M.R., Nejadhashemi, A.P., Daneshvar, F., Ross, D.M., Woznicki, S.A., Zhang, Z., Esfahanian, A.H., 2015. Optimization of conservation practice implementation strategies in the context of stream health. *Ecol. Eng.* 84, 1-12.
- Hilsenhoff, W.L., 1988. Rapid field assessment of organic pollution with a Family-Level Biotic Index. *J. N. Am. Benthol. Soc.* 7, 65–68.
- IPCC, 2013, What is a GCM?. URL: http://www.ipcc-data.org/guidelines/pages/gcm_guide.html (Last Accessed 11.15.16).
- Isaak, D.J., Wollrab, S., Horan, D., Chandler, G., 2012. Climate change effects on stream and river temperatures across the northwest U.S. from 1980–2009 and implications for salmonid fishes. *Clim. Chang.* 113, 499–524.
- Kerans, B.L., Karr, J.R., 1994. A Benthic Index of Biotic Integrity (B-IBI) for rivers of the Tennessee Valley. *Ecol. Appl.* 4, 768–785. 7
- Knight, J., Harrison, S., 2013. The impacts of climate change on terrestrial Earth surface systems. *Nature* 3, 24–29.
- Leigh, C., Stubbington, R., Sheldon, F., Boulton, A.J., 2013. Hyporheic invertebrates as bioindicators of ecological health in temporary rivers: A meta-analysis. *Ecol. Indic.* 32, 62–73.
- Lenat, D.R., 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. *J. N. Am. Benthol. Soc.* 7, 222–233.
- Lurgi, M., López, B.C., Montoya, J.M., 2012. Climate change impacts on body size and food web structure on mountain ecosystems. *Philos. T. Roy. Soc. B.* 367, 3050–3057.
- MDEQ, 1997. GLEAS Procedure #51 Survey protocols for wadeable rivers. Report No.: Fisheries Special. Report 25. Michigan Department of Environmental Quality, Surface Water Quality Division, Ann Arbor, MI.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *T. ASABE* 50, 885–900.
- NASS, 2014. USDA, National Agricultural Statistics Service, 2014 Michigan cropland data layer URL: http://www.nass.usda.gov/Research_and_Science/Cropland/metadata/metadata_mi14.htm (Last Accessed 2.15.16).
- NCDC, 2015. NOAA’s National Centers for Environmental Information (NCEI). URL: <http://www.ncdc.noaa.gov/data-access/land-based-station-data> (Last Accessed 2.15.16).
- NED, 2015. The National Map. URL: <http://nationalmap.gov/elevation.html> (Last Accessed 2.15.16).
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool theoretical documentation version 2009. Texas Water Resources Institute Technical Report No. 406.

- Nejadhashemi, A.P., Wardynski, B.J., Munoz, J.D., 2012. Large-scale hydrologic modeling of the Michigan and Wisconsin agricultural regions to study impacts of land use changes. *T. ASABE* 55, 821–838.
- NEX-GDDP, 2015. Advancing Research and Applications with NASA Climate Model Data. URL: <https://cds.nccs.nasa.gov/nex/> (Last Accessed 3.22.16).
- Pelletier, M.C., Gold, A.J., Gonzalez, L., Oviatt, C., 2012. Application of multiple index development approaches to benthic invertebrate data from the Virginian Biogeographic Province, USA. *Ecol. Indic.* 23, 176–188.
- Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The natural flow regime. *BioScience* 47, 769–784.
- Runkel, R.L., Crawford, C.G., Cohn, T.A., 2004. Load estimator (LOADEST): a FORTRAN program for estimating constituent loads in streams and rivers (USGS Numbered Series No. 4-A5), Techniques and Methods.
- Schmatz, D.R., Luterbacher, J., Zimmermann, N.E., Pearman, P.B., 2015. Gridded climate data from 5 GCMs of the Last Glacial Maximum downscaled to 30 arc s for Europe. *Clim. Past* 11, 2585–2613.
- SSURGO, 2014. SSURGO Data Downloader (beta). URL: <https://www.arcgis.com/home/item.html?id=a23eb436f6ec4ad6982000dbaddea5ea> (Last Accessed 2.15.16)
- Thrasher, B., Xiong, J., Wang, W., Melton, F., Michaelis, A., Nemani, R., 2013. Downscaled climate projections suitable for resource management. *Eos T. AGU* 94, 321–323.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. The river continuum concept. *Can. J. Fish. Aquat. Sci.* 37, 130–137.
- Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O., Bairlein, F., 2002. Ecological responses to recent climate change. *Nature* 416, 389–395.
- Weibull, W., 1939. The statistical theory of the strength of materials. *Roy. Swedish Acad. Eng. Sci.* 151, 1–45.
- Woznicki, S.A., Nejadhashemi, A.P., 2014. Assessing uncertainty in best management practice effectiveness under future climate scenarios. *Hydrol. Process.* 28, 2550–2566.
- Woznicki, S.A., Nejadhashemi, A.P., Abouali, M., Herman, M.R., Esfahanian, E., Hamaamin, Y.A., Zhang, Z., 2016a. Ecohydrological modeling for large-scale environmental impact assessment. *Sci. Total Environ.* 543, Part A, 274–286.
- Woznicki, S. A., Nejadhashemi, A. P., Tang, Y., Wang, L. 2016b, Large-scale climate change vulnerability assessment of stream health. *Ecol. Indic.* 69, 578-594.
- Woznicki, S.A., Nejadhashemi, A.P., Ross, D.M., Zhang, Z., Wang, L., Esfahanian, A.-H., 2015. Ecohydrological model parameter selection for stream health evaluation. *Sci. Total Environ.* 511, 341–353.

Zhu, K., Woodall, C.W., Ghosh, S., Gelfand, A.E., Clark, J.S., 2014. Dual impacts of climate change: forest migration and turnover through life history. *Glob. Change Biol.* 20, 251–264.

Table 1. Summary of the near future climate conditions (2040-2060) for the 16 weather stations

Model name	RCP 4.5			RCP 8.5		
	Precipitation (mm)	Minimum Temperature (°C)	Maximum Temperature (°C)	Precipitation (mm)	Minimum Temperature (°C)	Maximum Temperature (°C)
ACCESS1-0	778 - 878	3.93 - 6.19	15.24 - 17.29	676 - 822	4.94 - 7.10	16.36 - 18.43
BCC-CSM1-1	757 - 851	3.79 - 6.06	14.82 - 16.79	850 - 946	4.41 - 6.63	15.25 - 17.17
BNU-ESM	843 - 941	4.47 - 6.75	15.39 - 17.23	853 - 955	4.97 - 7.12	15.69 - 17.53
CanESM2	831 - 908	4.21 - 6.52	15.08 - 17.02	798 - 880	4.77 - 7.14	15.74 - 17.76
CCSM4	742 - 814	4.14 - 6.30	15.44 - 17.46	817 - 916	4.36 - 6.35	15.56 - 17.57
CESM1-BGC	835 - 917	3.71 - 5.68	15.02 - 16.96	850 - 964	4.45 - 6.48	15.46 - 17.39
CNRM-CM5	862 - 990	3.68 - 5.90	14.18 - 16.01	796 - 934	4.49 - 6.80	15.11 - 17.05
CSIRO-MK3-6-0	901 - 1007	3.58 - 6.15	14.25 - 16.46	872 - 974	3.89 - 6.37	14.50 - 16.66
GFDL-CM3	940 - 1051	4.41 - 6.61	15.47 - 17.36	939 - 1074	5.12 - 7.19	16.00 - 17.80
GFDL-ESM2G	894 - 1001	2.97 - 5.35	13.99 - 16.01	923 - 1037	3.37 - 5.80	14.21 - 16.27
GFDL-ESM2M	813 - 905	2.92 - 5.21	13.80 - 15.76	781 - 881	3.32 - 5.52	14.26-16.21
INMCM4	769 - 864	2.84 - 5.14	13.59 - 15.51	744 - 834	3.01 - 5.36	13.85 - 15.89
IPSL-CM5A-LR	785 - 897	3.69 - 6.07	14.60 - 16.69	805 - 904	4.49 - 6.85	15.29 - 17.38
IPSL-CM5A-MR	772 - 863	3.75 - 6.12	14.64 - 16.75	749 - 849	4.55 - 6.89	15.31 - 17.44
MIROC-ESM	820 - 915	5.05 - 7.12	15.97 - 17.63	833 - 939	6.29 - 8.21	17.30 - 18.94
MIROC-ESM-CHEM	868 - 970	4.98 - 7.17	15.92 - 17.77	852 - 943	6.04 - 8.05	16.90 - 18.70
MIROC5	790 - 924	3.98 - 6.29	15.51 - 17.64	861 - 962	4.76 - 7.00	16.34 - 18.32
MPI-ESM-LR	778 - 874	3.52 - 5.77	14.26 - 16.19	790 - 890	4.46 - 6.82	15.28 - 17.38
MPI-ESM-MR	875 - 969	3.45 - 5.78	14.08 - 16.01	833 - 934	4.41 - 6.67	14.92 - 16.87
MRI-CGCM3	817 - 948	3.09 - 5.36	13.73 - 15.51	912 - 1046	3.53 - 5.85	14.04 - 15.84
NorESM1-M	755 - 840	3.74 - 5.89	15.02 - 16.96	823 - 924	4.29 - 6.40	15.46 - 17.44

Table 2. Summary of SWAT model performance for streamflow

USGS Station No.	Calibration (2001-2005)			Validation (2006-2010)		
	NSE	PBIAS (%)	RSR	NSE	PBIAS (%)	RSR
04144500	0.64	20.88	0.6	0.62	9.12	0.62
04147500	0.61	9.43	0.62	0.65	-10.04	0.59
04148140	0.54	9.69	0.68	0.53	-24.98	0.68
04148500	0.65	26.3	0.59	0.76	8.08	0.49
04151500	0.59	15.14	0.64	0.68	12.62	0.57
04154000	0.54	5.51	0.68	0.54	13.24	0.68
04155500	0.63	20.1	0.61	0.59	1.29	0.64
04156000	0.73	3.27	0.52	0.74	8.98	0.51
04157005	0.79	11.66	0.46	0.8	10.3	0.44

Table 3. Summary of SWAT model performance for water quality parameters

Parameter	Calibration (1998-2001)			Validation (2002-2005)		
	NSE	PBIAS(%)	RSR	NSE	PBIAS(%)	RSR
Sediment	0.59	-1.85	0.64	0.57	28.42	0.66
Total Nitrogen	0.52	28.16	0.69	0.58	41.3	0.64
Total Phosphorous	0.53	5.27	0.68	0.81	14.16	0.43

Table 4. Summary of two-phase model performances

Biotic index	Stream Group					
	All		Headwater		Medium-sized and large	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
HBI	0.83	0.27	0.91	0.21	0.97	0.11
FIBI	0.70	4.14	0.94	1.71	0.94	2.01
EPT	0.72	2.04	0.97	0.73	0.91	1.10

Figure Captions:

Figure 1. The study area

Figure 2. Climate stations and macroinvertebrates health monitoring sites

Figure 3. Modeling process

Figure 4. a) Current overall stream health condition b) dominant land uses

Figure 5. Probability of decline in overall stream health under future climate scenarios

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 4 - colour

Figure 5 - colour