Dan Li
Dan Myers
Data Visualization
9/26/2019

Climate Change and Great Lakes Tributaries

Project Proposal

Daniel T. Myers[a], Dan Li[b]

[a] Department of Geography, Indiana University, Student Building 120, 701 E. Kirkwood Avenue, Bloomington, IN 47405, USA (Email: myersdt@iu.edu)

[b] Department of History and Philosophy of Science and Medicine, Indiana University, Morrison Hall 308, 1165 E 3rd St, Bloomington, IN 47405, USA (Email: dli1@iu.edu)

Abstract

Climate change is impacting rivers of the Great Lakes Basin, creating risks for ecosystems, communities, and economies. Existing visualizations of climate change impacts often display data on maps or time-series charts; however, they do not convey clearly what is happening in the Basin. An Indiana University research project is modeling these impacts using a hydrologic model and an ensemble of climate projections. We are using data from this study to create three visualizations so that the general public can understand the impacts and have confidence in the model. The first visualizations is a data map that shows our confidence in the model at 190 calibration stations along rivers in the Basin. The second visualization is a series of time-series charts that show how streamflow, water temperature, and sediment loading rates for these rivers will be affected by climate change during the 21st century. The third visualization is another map that shows spatially (using a data map) how hydrology is projected to change throughout the basin during the 21st century. These visualizations are vital for communicating the science to the general public.

Introduction

Climate change is impacting the Earth's rivers by altering the hydrologic cycle, which determines the partitioning and timing of available water (Jiménez Cisneros et al., 2015; Knouft and Ficklin, 2017). These impacts are witnessed in historical and simulated changes to streamflow, water temperature, and sediment loading, posing risks for the ecosystems, communities, and economies that depend on these rivers. Ecologically, these impacts are causing a loss of river biodiversity, fisheries, and habitat (Ficke et al., 2007; Knouft and Ficklin, 2017). Socially, these impacts are causing water shortages for communities, producing extreme floods, and creating problems for water quality and treatment (Ficke et al., 2007; Jiménez Cisneros et al., 2015; Kundzewicz et al., 2008). Economically, these impacts cause repercussions for navigation and commercial transportation, hydroelectric power generation, invasive species control, and harbors and marinas (Campbell et al., 2015; Environment Canada and USEPA, 1995).

Rivers of North America's Great Lakes Basin, the Earth's largest system of surface freshwater, are vulnerable to these impacts (Environment Canada and USEPA, 1995; Figure 1). Average annual air temperatures in the Basin are rising while precipitation patterns are projected to shift seasonally, with more precipitation in winter and spring (Hayhoe et al., 2010; Byun et al., 2019). These climate changes are impacting the hydrology of the Basin's rivers and streams (Bartolai et al., 2015).

Figure 1. Study area map of the Great Lakes Basin (USACE, https://commons.wikimedia.org/wiki/File:Great_lakes_basin.jpg)

<u>Importance of model robustness for credibility and confidence</u>

Due to the ecological, social, and economic impacts of runoff affected by climate change, the scientific research on these impacts may be of interest to many non-scientists, or at least people unfamiliar with climate models. Climate model projections come with a host of uncertainties for various reasons: for example, there are various ways to represent aspects of the climate system, the resolution of climate models is too coarse to capture some processes which then have to be parameterized, and the effects of some feedbacks (e.g. water vapor feedbacks) are different in different climate models. One way to address this uncertainty is to evaluate the robustness of climate models (see below). A further challenge is to communicate this robustness to interested parties and help them understand how uncertainty relates to robustness. We devise a way to visualize robustness in an effort to make this communication more effective. If a non-scientist can see how climate models are robust without learning the technical details of the models or the climate system, then they can be more easily informed and make better decisions in their respective fields.

Climate models are incredibly complex. In order to compare the prediction results and utility of models, a frequently sought feature of models among scientists is “robustness”, although the term “robustness” is used differently by different authors (e.g. Woodward (2006), Weisberg and Reisman (2008), Schupbach (2018)). One strategy of evaluating the robustness of models, known as “robustness analysis”, boils down to measuring the invariance or agreement of results produced by models with possibly different initial assumptions and inputs. In other words, if a result R occurs respectively under the same set of core assumptions C and different auxiliary assumptions A_1,A_2,…, then we may have the confidence to conclude that the core C is driving the result. Thus, any reasonably constructed models with this core C, which have been shown to reliably give some result R despite differing auxiliary assumptions, are considered more robust than those equally reasonably constructed models that lack such a core C.

However, in practice, complex models are not Lego blocks. They cannot be broken into isolated units and swapped with different parts without causing changes in other parts. This makes “robustness” incredibly hard to track. Some scholars claim that the introduction of auxiliary assumption would alter the structure of models and possibly the core assumption, making it impossible to track “what… is being compared across cases” (Lisciandra, 2017). Some other notion of robustness, in particular, “model robustness” by Lloyd (2015) was invented to mitigate this holism issue by adding considerations of the strengths of independent evidential support for both the core assumption and the various auxiliary assumptions.

<u>Robustness and uncertainty</u>

The concept of “robustness” of models is essentially aimed as a management for uncertainty. Climate model projections come with a host of uncertainties for various reasons: for example, there are various ways to represent aspects of the climate system, the resolution of climate models is too coarse to capture some processes which then have to be parameterized, and the effects of some feedbacks (e.g. water vapor feedbacks) are different in different climate models.

One way to address this uncertainty is to evaluate the robustness of climate models. If we are uncertain about the projection of a single model, perhaps we can reduce this uncertainty a bit by looking at the projection of another model, which has at least some different assumptions but which shares the same key causal features that are well-supported by physical theory and empirical evidence. If the second model agrees with the first model about some variable of interest, we may be a bit more confident that the model projection is not merely a product of its simplifying assumptions (parameterizations) but is rather due to the core features of the model, including the radiative forcing, the basic physics equations in the model, and the empirically well-supported assumptions the model makes. The basic idea is that if both models share these latter features but have different simplifying assumptions, we may have more confidence in their projections.

Existing Visualizations

Map-based visualization for climate change impact

Existing visualizations for climate change and its impacts often display data on maps or time series charts. Map-based visualizations from existing research show a slight increase in runoff of the Great Lakes Basin by the end of the 21st century at 5% above 1981-2000 levels (Kundzewicz et al., 2007; Figure 2). The existing research, however, uses global climate models with coarse spatial resolution, resulting in a scale mismatch in existing visualizations that makes them not applicable to hydrology studies of individual watersheds (Kundzewicz et al., 2007). This scale mismatch can be seen in the pixelated displays of data on the map. Additionally, the color scale of the Kundzewicz et al. map shows clearly that redder areas will experience droughts and bluer areas will experience more precipitation. The scale of these color changes has been hand-selected however, binning the changes at 5% intervals at first, then increasing the interval to 10% and even 20% intervals for higher percents. This hand-selection does not visualize areas with zero change; thus, it appears that the amount of runoff for the whole world will be changing, where in reality areas between -5% and 5% are not changing much and should be represented by a neutral color. Thus, this visualization is insufficient for showing the impacts of climate change in the Great Lakes Basin.

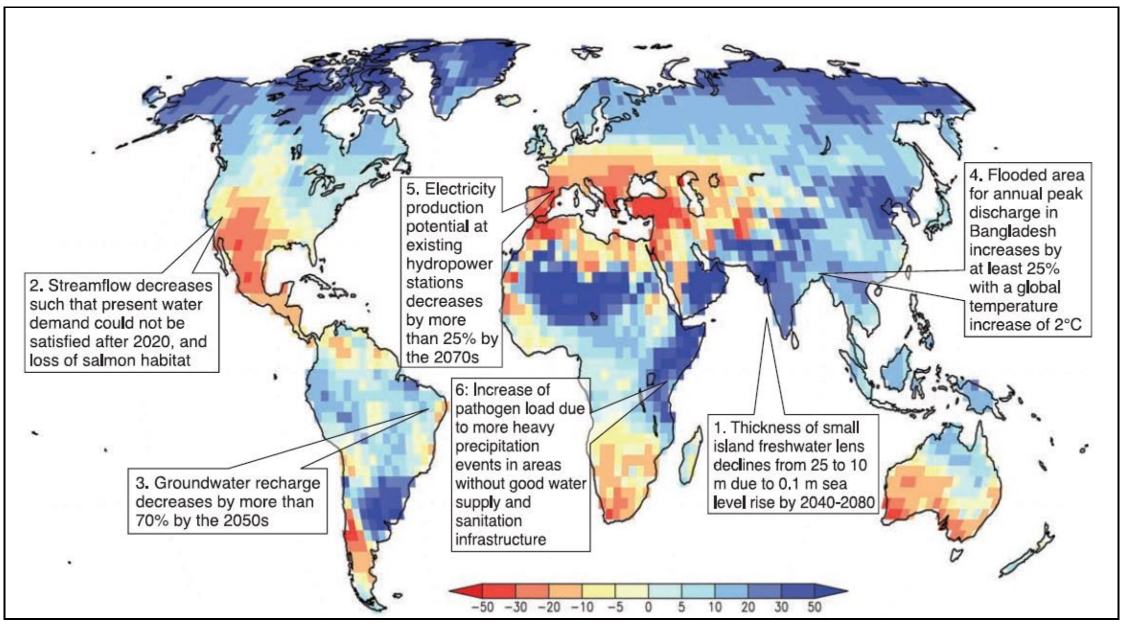

Figure 2. Mean projected change in annual runoff (%) between historic (1981-2000) and future (2081-2100) periods (Kundzewicz et al. 2007).

Time-series charts for model projections

Existing visualizations for time-series charts commonly show several models or datasets to understand both the general trends in the data and the uncertainty. An example is this visualization of sea level rise over time from the Intergovernmental Panel on Climate Change's 5th Assessment Report (IPCC, 2014; Figure 3). The different colors and widths of lines represent different sources of global sea level data that were used to make the analysis, while uncertainty is shown by the variations in the lines. However, all these different lines make the figure confusing: front-drawn lines obscure the lines behind them and the uncertainty could be made more clear. One way to improve this is to make all the lines the same color but semi-transparent; thus, where there is strong agreement between lines, there will be overlap of the semi-transparent colors and a darker overall color. Uncertainty would be displayed by lighter colors, due to the semi-transparency. Additionally, the chart has much blank space between the years 1850 and 1900 where there is insufficient data. This leaves a viewer asking: “Is there a cyclical or decreasing trend in the data here that the designers want to cover up?” Setting the lower x-axis limit to 1900 would remove this confusing blank space; however, it would also make the slope of the sea level change smaller, which could be interpreted incorrectly as an insignificant change by a viewer. Thus, the entire chart should be shortened longitudinally to accommodate this. A time-series visualization for the Great Lakes Basin should incorporate these improvements to more-effectively communicate the hydrological trends and uncertainties.

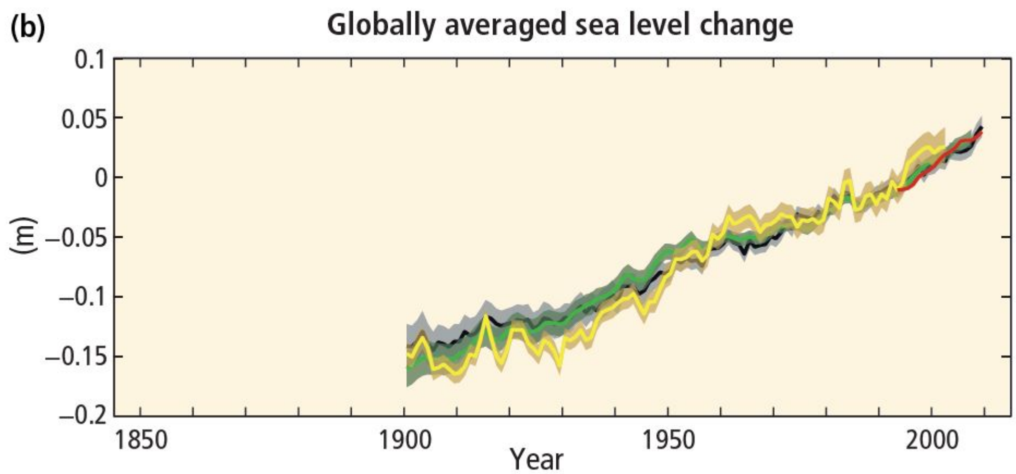

Figure 3. Globally averaged change in sea levels from 1900 to 2010. Different lines represent different data sources (IPCC, 2014).

<u>Visualizations for the reliability of models tuning</u>

Existing visualizations of the reliability of hydrologic models often display statistics that represent the fit between simulated time-series and historic observations. These can be plotted on a map to visualize how effectively the model works at the locations with historic data (Neupane et al. 2018; Figure 4). A statistic such as Willmott’s d_r is commonly used (Willmott et al. 2012). In this case, a d_r approaching 1.0 is excellent, while below 0.65 is less reliable. The color scheme of green, yellow, and grey clearly shows the reliability of the model at each location when printed in color. However, in greyscale, the colors are indistinguishable. A visualization of the reliability of the Great Lakes Basin hydrologic model will use a color scheme with varying levels of lightness that are distinguishable if printed in greyscale.

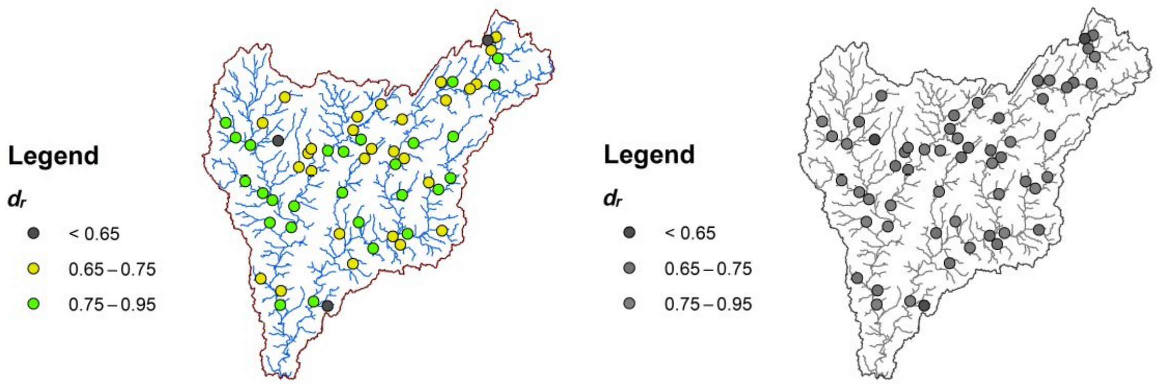

Figure 4. Reliability of a hydrologic model at historic observation stations in the Mobile River, Alabama visualized in color (left) and greyscale (right) (Neupane et al. 2018).

Existing visualizations and methods for uncertainty

There is a significant amount of research on uncertainty analysis and visualization with a variety of uncertainty metrics defined and reframed. Broadly, there are three major expressions of uncertainty, portrayed by data divergence, the estimation results of unknown parameters, and the fuzziness in clustering or classifying data. (Liang, 2016) Data divergences have to do with the variability of simulation results from multiple runs or different model settings. The estimation of unknown parameters has to do with confidence interval or posterior probabilistic distribution in Bayesian statistics. The clustering of data as a characterization of uncertainty is defined as the fuzziness among different membership functions.

The exploration of statistical presentation of robustness has become an emerging research topic as well. For example, Knutti and Sedlacek (2013) defines robustness as the ratio of model spread to the predicted change, that is, R = 1- A1/A2, where A1 is the integral of the squared area between two cumulative density functions characterizing the individual model projections and the multi-model mean projection and A2 is the integral of the squared area between two cumulative density functions characterizing the multi-model projection and the historical climate.

For the scope of this project, we favor the approach of descriptive statistics for its simplicity and effective visualization. For example, Knutti and Sedlacek (2013) uses mean and standard deviation to represent model uncertainty. In the following graph, they use four colored lines to present the means of model family predictions under four different scenarios (and historical data) and one standard deviations of the predictions under each scenario as shading. On the right side, they also use error bars to represent the uncertainty between the year 2080-2099.

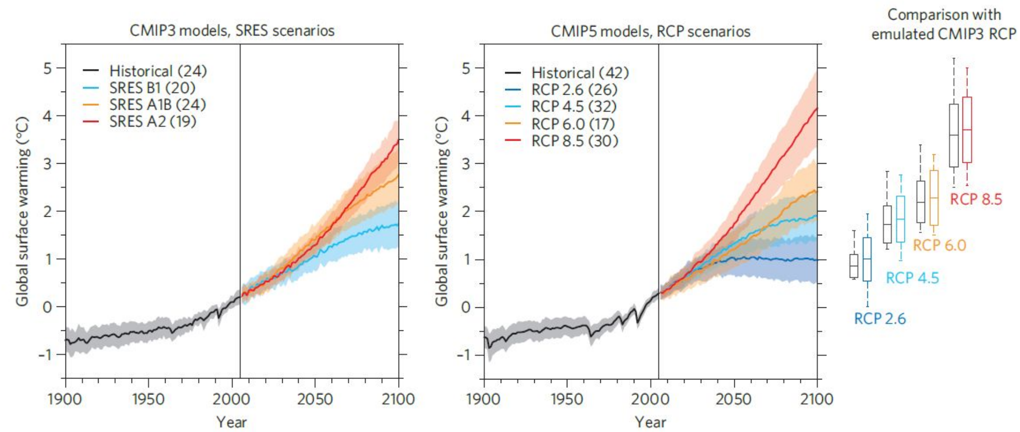

Figure 5. Global temperature change and uncertainty. Global temperature change (mean and one standard deviation as shading) relative to 1986-2005 for the SRES scenarios run by CMIP3 and the RCP scenarios run by CMIP5. The number of models is given in brackets. The box plots on the right are given for 2080-2099 for CMIP5 (colors) and for the MAGICC model calibrated to 19 CMIP3 models (black), both running the RCP scenarios. (Kunutti and Sedlack, 2013)

This is a very straight-forward and relatively easy to implement technique for visualization climate model uncertainty where you run a family of models under some scenarios and see how variant the projections are.

With regard to robustness visualization, Knutti and Sedlacek (2013), as mentioned, defines a robustness ration R, which is essentially how spread the models projections are. They select the earth surface with high robustness (black, R>0.8) and shows this fraction of land surface in relation to the global surface warming under different models.

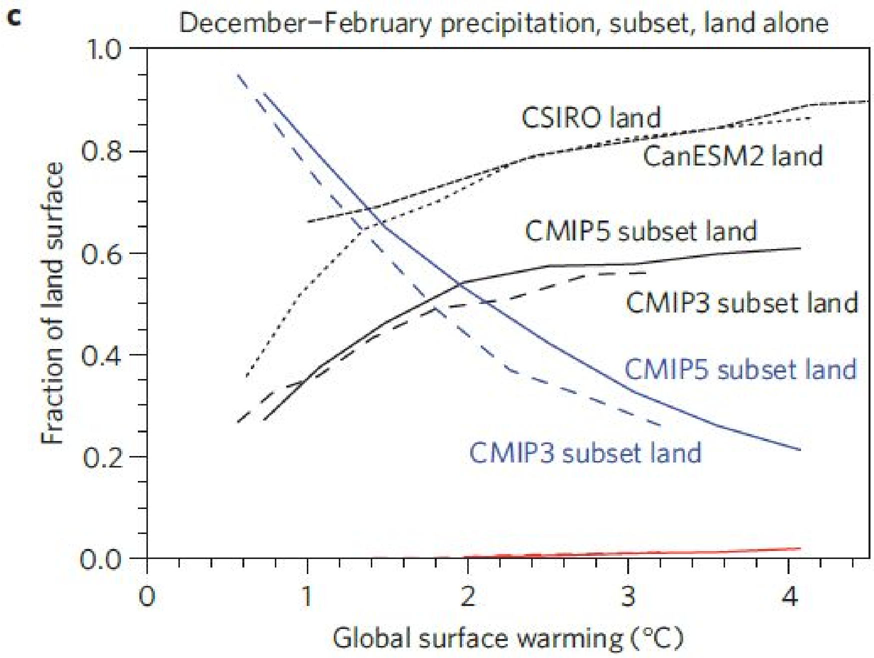

Figure 6. Model robustness for precipitation for land. Fraction of the Earth's land surface with high robustness (black, R>0.8), no significant change (blue) and inconsistent model responses (red) illustrated for December-February precipitation change for CMIP3 and CMIP5. Fractions are shown as a function of global temperature change rather than time, which largely eliminates the differences in scenarios. (Knutti and Sedlack, 2013)

Although this graph does show that some models are more robust when global surface warming is higher, in the sense that the projections of more global surface are with higher robustness, it would be interesting to see exactly the projection of which part of the global surface is with more robustness. Admittedly, this visualization of robustness has a different definition than the ones previously mentioned, but we have not concluded a better way to represent mathematically the concept of “model robustness” with additional empirical evidence support taken into consideration as a component.

Objectives

It is important to understand climate change impacts at a local scale in order to sustainably manage rivers and prepare for these risks, both within the Great Lakes Basin and around the world. It is also vital to be able to present these understandings to the general public and stakeholders in a way that is intuitive, aesthetically pleasing, and easy to comprehend. Our project uses novel data visualization techniques to communicate clearly and effectively the impacts of climate change on Great Lakes Tributaries. Our objectives are to:

1. Create visualizations to show projected impacts to rivers and streams of the Great Lakes Basin throughout the 21st century. This includes time series charts and maps that visualize spatial and temporal climate change impacts.
2. Visualize uncertainties in hydrologic projections due to the use of 20 different options for climate models from the Climate Model Intercomparison Project Phase 5 ensemble.
3. Explore the visualization techniques for robustness analysis and the concept of model robustness.

Methods

The datasets for these visualizations will be from the research project "Hydrologic Modeling for Sustainable Management of Great Lakes Basin Rivers Impacted by Climate Change" funded by the IU Sustainability Research Development Grant. This project forces a hydrology model with an ensemble of downscaled and bias-corrected global climate model outputs. The project data includes model confidence data (validation results for 190 river stations with historical observations) and monthly hydrologic projections from 1950-2099. Although the model for this project is not finalized yet, we have preliminary data that will be used until the final projections become available later this year.

For the model reliability visualizations, each historical observation station (n=190) has data for latitude, longitude, and validation statistics (Table 1). An example validation statistic we use is the Willmott's d_r, which is on a scale from -1 (poor calibration) to 1 (excellent calibration) (Willmott et al. 2012). We will use the latitude and longitude data to place these river stations on a map of the Great Lakes Basin and use either colors or symbols to show how well our model is performing against observed data. ArcGIS or Jupyter Notebook with Geopandas will be used to make these maps.

Table 1. Head of robustness data showing geographic coordinates and validation with Willmott's dr statistic.

Calibration Station	Latitude	Longitude	Streamflow (d_r)
1	49.5979922564	-87.9642156679	0.39
2	48.9260761327	-87.6899350553	0.51
3	48.9047240215	-88.3759454033	0.41
4	48.8480328924	-86.6069774735	0.46
5	48.8227860408	-88.5319012409	0.51
6	48.7781993661	-86.8839475334	0.47
7	48.7728205966	-86.2936593433	0.48
8	48.6878865469	-86.2129348282	0.52
9	48.5309431581	-89.5939070020	0.35
10	48.2918848282	-89.8084367575	0.50

The time series data for hydrological projections will include monthly projections of streamflow, water temperature, and suspended sediment load for the Basin's rivers from 1950-2099 (Table 2). There will be one set of projections for each climate model in the ensemble (n=20). We will calculate the monthly mean projection by averaging the ensemble results, then visualize the uncertainty of the models by plotting the errors of the projections. We will then make a time series chart showing how climate change is affecting the hydrology of the basin over the 1950-2099 time period (n=1800 months). We will work with the axes, colors, and other chart features to ensure that the visualization is appropriate for its context. Jupyter Notebook with matplotlib and pandas will be used to make this visual. We will also use python viridis package for proper color scheme. We will also display changes over time using points on a map, similar to the calibration stations map. For example, the color or size of the points could symbolize percent change in streamflow from the present to 2099 at each calibration station. These visualizations will improve upon the existing visualizations by incorporating the critiques discussed.

Table 2. Head of time series data for historically-observed and simulated streamflow for Station #167.

Year	Month	Observed Streamflow (cms)	Simulated Streamflow (cms)
1950	1	52.8	15.2
1950	2	44.2	26.4
1950	3	54.8	23.5
1950	4	68.6	29.0
1950	5	22.0	8.4
1950	6	16.2	9.0
1950	7	20.3	12.7
1950	8	13.7	9.4
1950	9	15.7	12.8
1950	10	9.6	6.2

Expected Outcomes

1. We will have a map with the reliability of this hydrological model over 190 river stations in the Great Lakes Basin area, similar to Figure 4.

2. We will have a series of visualizations for monthly projections of streamflow, water temperature, and suspended sediment load respectively for the Basin's rivers, using line chart with shadings as uncertainties, similar to Figure 5 (left).

3. We will have a data map that visualizes the hydrology model projections over the century.

4. We will use a box chart to visualization the overall uncertainties regarding our projections of streamflow, water temperature, and suspended sediment load, similar to Figure 5 (right).

References/Acknowledgements

Bartolai, A.M., He, L., Hurst, A.E., Mortsch, L., Paehlke, R., Scavia, D., 2015. Climate change as a driver of change in the Great Lakes St. Lawrence River Basin. Journal of Great Lakes Research 41, 45–58. https://doi.org/10.1016/j.jglr.2014.11.012

Byun, K., Chiu, C.M., Hamlet, A.F., 2019. Effects of 21st century climate change on seasonal flow regimes and hydrologic extremes over the Midwest and Great Lakes region of the US. Science of the Total Environment 650, 1261–1277. https://doi.org/10.1016/j.scitotenv.2018.09.063

Campbell, M., Cooper, M.J., Friedman, K., Anderson, W.P., 2015. The economy as a driver of change in the Great Lakes-St. Lawrence River basin. Journal of Great Lakes Research 41, 69–83. https://doi.org/10.1016/j.jglr.2014.11.016

Environment Canada, USEPA, 1995. The Great Lakes: An Environmental Atlas and Resource Book, Cartography. https://doi.org/10.1038/184665a0

Ficke, A.D., Myrick, C.A., Hansen, L.J., 2007. Potential impacts of global climate change on freshwater fisheries. Reviews in Fish Biology and Fisheries 17, 581–613. https://doi.org/10.1007/s11160-007-9059-5

Hayhoe, K., VanDorn, J., Croley, T., Schlegal, N., Wuebbles, D., 2010. Regional climate change projections for Chicago and the US Great Lakes. Journal of Great Lakes Research 36, 7–21. https://doi.org/10.1016/j.jglr.2010.03.012

IPCC, 2014. Climate Change 2014 Synthesis Report: Summary for Policymakers, in: Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. p. 32 pages. https://doi.org/10.1017/CBO9781107415324

Jiménez Cisneros, B.E., Oki, T., Arnell, N.W., Benito, G., Cogley, J.G., Döll, P., Jiang, T., Mwakalila, S.S., Kundzewicz, Z., Nishijima, A., 2015. Freshwater resources, in: Climate Change 2014 Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects. p. 41 pages. https://doi.org/10.1017/CBO9781107415379.008

Knouft, J.H., Ficklin, D.L., 2017. The Potential Impacts of Climate Change on Biodiversity in Flowing Freshwater Systems. Annual Review of Ecology, Evolution, and Systematics 48, 111–133. https://doi.org/10.1146/annurev-ecolsys-110316-022803

Knutti, Reto, and Jan Sedláček. “Robustness and Uncertainties in the New CMIP5 Climate Model Projections.” Nature Climate Change 3, no. 4 (April 2013): 369–73. https://doi.org/10.1038/nclimate1716.

Kundzewicz, Z.W., Mata, L.J., Arnell, N.W., Doll, P., Jimenez, B., Miller, K., Oki, T., Sen, Z., Shiklomanov, I., 2008. The implications of projected climate change for freshwater resources and their management. Hydrological Sciences Journal 53, 3–10. https://doi.org/10.1623/hysj.53.1.3

Kundzewicz, Z.W., Mata, L.J., Arnell, N.W., Döll, P., Kabat, B., Jimenez, B., Miller, K. a, Oki, T., Sen, Z., Shiklomanov, I., 2007. Freshwater resources and their management, in: Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/CBO9781107415324.004

Liang, Xing. “Visualizing Numerical Uncertainty in Climate Ensembles.” Masters of Science, Arizona State University, 2016.

Lisciandra, Chiara. “Robustness Analysis and Tractability in Modeling.” European Journal for Philosophy of Science 7, no. 1 (January 1, 2017): 79–95. https://doi.org/10.1007/s13194-016-0146-0.

Lloyd, Elisabeth A. “Model Robustness as a Confirmatory Virtue: The Case of Climate Science.” Studies in History and Philosophy of Science Part A 49 (February 1, 2015): 58–68. https://doi.org/10.1016/j.shpsa.2014.12.002.

Neupane, R.P., Ficklin, D.L., Knouft, J.H., Ehsani, N., Cibin, R., 2018. Hydrologic responses to projected climate change in ecologically diverse watersheds of the Gulf Coast, United States. International Journal of Climatology 39, 2227–2243. https://doi.org/10.1002/joc.5947

Schupbach, Jonah N. “Robustness Analysis as Explanatory Reasoning.” The British Journal for the Philosophy of Science 69, no. 1 (March 1, 2018): 275–300. https://doi.org/10.1093/bjps/axw008.

Weisberg, Michael, and Kenneth Reisman. “The Robust Volterra Principle.” Philosophy of Science 75, no. 1 (January 1, 2008): 106–31. https://doi.org/10.1086/588395.

Willmott, C.J., Robeson, S.M., Matsuura, K., 2012. Short Communication A refined index of model performance. International Journal of Climatology 33, 1053–1056. https://doi.org/10.1002/joc.2419

Woodward, Jim. “Some Varieties of Robustness.” Journal of Economic Methodology 13, no. 2 (June 1, 2006): 219–40. https://doi.org/10.1080/13501780600733376.