Climate Change and Great Lakes Tributaries

Final Report

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

INFO I-590 Data Visualization

Fall 2019

[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)

Table of Contents

Abstract

Climate change is impacting rivers of the Great Lakes Basin, creating risks for ecosystems, communities, and economies. Effective data visualizations are imperative for communicating the science of these impacts to the public and decision makers. However, existing visualizations of climate change impacts often do not convey clearly the climate science to the general public. Additionally, effective data visualizations for model robustness and uncertainty are needed to communicate these important aspects of climate science. An Indiana University research project is modeling climate change impacts to the Great Lakes Basin using a hydrologic model and an ensemble of climate projections. We used data from this study to create time series visualizations of climate change impacts and uncertainties to show robustness, data maps of impacts shown spatially across the Basin, and a data map of model tuning evaluation. These visualizations improve upon existing techniques for displaying climate change impacts and model robustness, effectively communicating the science to the general public.

Introduction

Motivation

Caused by increases in greenhouse gas emissions, climate change has been impacting ecosystems and communities across all continents and oceans, leading to warm temperature extremes, heavy precipitation events, and an increase in sea levels (IPCC 2014). These climatic changes are affecting the Earth's freshwater resources by altering the hydrologic cycle, which determines the partitioning and timing of available water (Jiménez Cisneros et al., 2015; Knouft and Ficklin, 2017).

Several styles of visualizations for climate change impacts currently exist, including 1) time series visualizations of impacts and uncertainty, 2) map-based visualizations of impacts, and 3) visualizations for model tuning, and 4) visualizations of model robustness. However, a major difficulty with communicating science regarding climate change impacts to freshwater is that the physical interactions these visualizations represent are embedded in complicated models and charts, which are difficult for the public to understand. Creating effective data visualizations for this purpose is demanding for time and effort, while also involving knowledge of best practices (Dasgupta et al. 2015). Without proper visualizations, it can be difficult to educate the public about climate science, causing incorrect interpretations that ultimately could lead to misguided policies.

The communication of model robustness is particularly important for providing credibility and confidence in climate science to the general public. However, model robustness is frequently absent from the communication of climate science, which often focuses more on the end-result (predictions) and the expected impacts to society. Effective visualizations of model robustness are needed to improve this communication.

Our study aimed to create effective data visualizations for climate change impacts to rivers in the Great Lakes Basin, as well as communicating the robustness of the underlying models that simulate these impacts. To do this, we used data from an ongoing Indiana University study and a suite of data visualization best practices including choice of data structure, visualization type, color preference, scale, data-ink ratio, and cartographic display. The visualizations we created are of utmost importance for communicating climate change impacts to the general public, in order to sustainably manage rivers and prepare for climate change risks, both within the Great Lakes Basin and around the world. The visualizations are also vital for communicating model robustness in a way that is intuitive, aesthetically pleasing, and easy to comprehend.

Background

Climate Change Impacts to Great Lakes Tributaries

Climate change's effects on freshwater systems have been witnessed in historical and simulated changes to streamflow, water temperature, and sediment loading. These changes pose risks for the ecosystems, communities, and economies that depend on the freshwater resources. 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)

Importance of model robustness for credibility and confidence

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 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 $A_1, A_2, \ldots$, 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.

Robustness and uncertainty

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

Time-series visualizations 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.

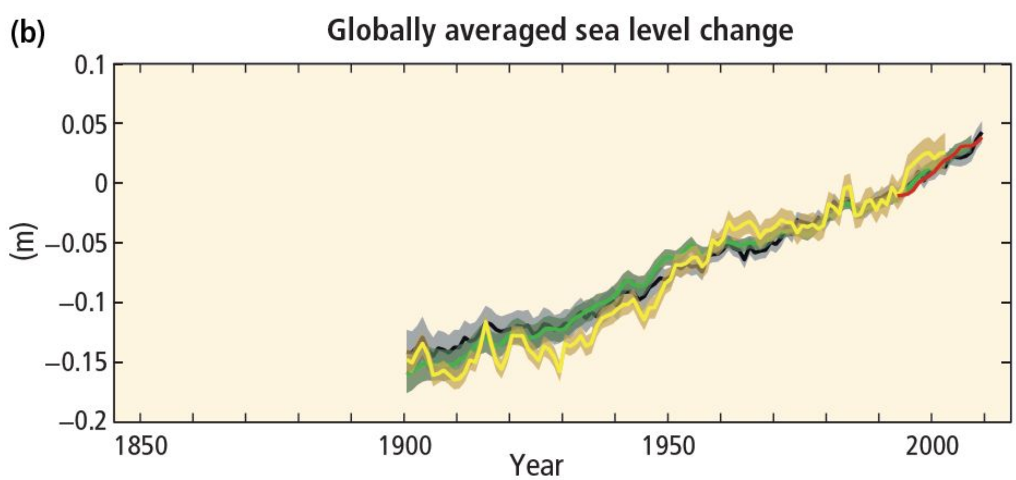

Figure 2. Globally averaged change in sea levels from 1900 to 2010. Different lines represent different data sources (IPCC, 2014).

Map-based visualizations 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.

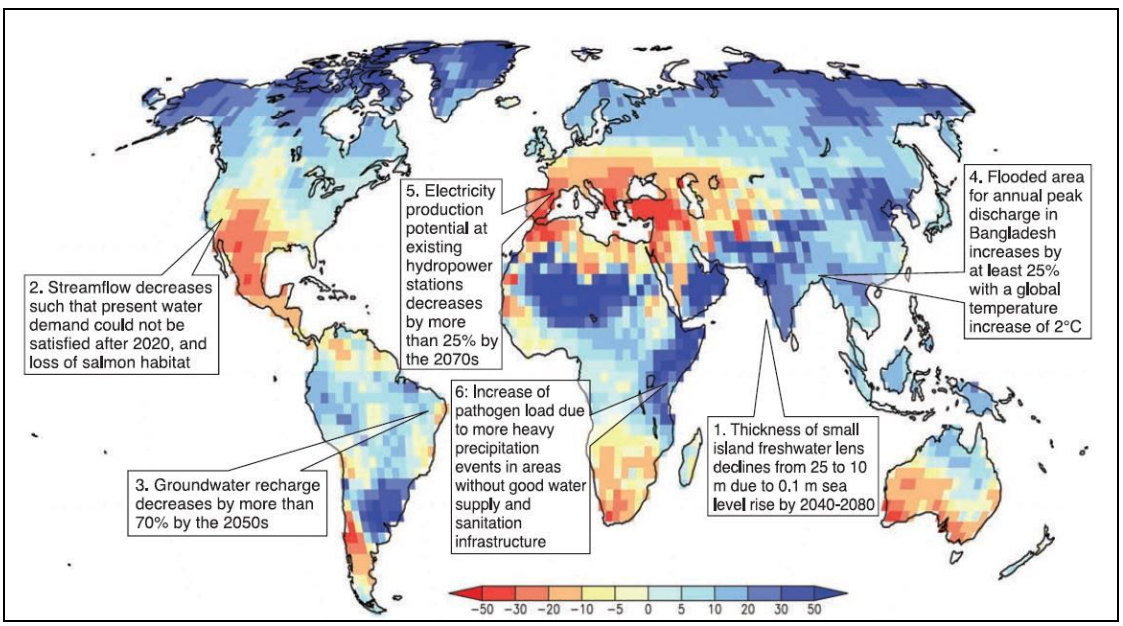

Figure 3. Mean projected change in annual runoff (%) between historic (1981-2000) and future (2081-2100) periods (Kundzewicz et al. 2007).

Visualizations for the reliability of models tuning

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), which is on a scale from -1.0 (awful) to 1.0 (excellent). 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.

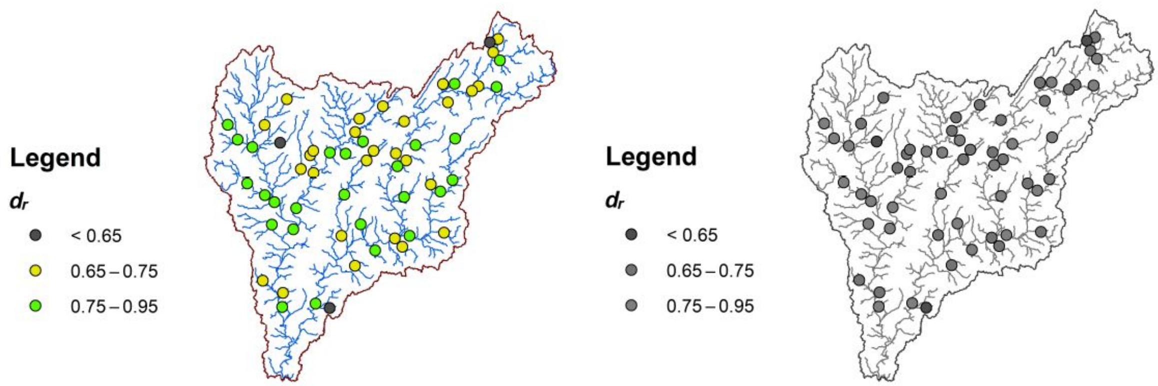

Figure 4. Reliability of a hydrologic model at historic observation stations in the Mobile River, Alabama visualized in color (left) and greyscale (right) (Neupane et al. 2018).

Existing visualizations and methods for model robustness and 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) use 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.

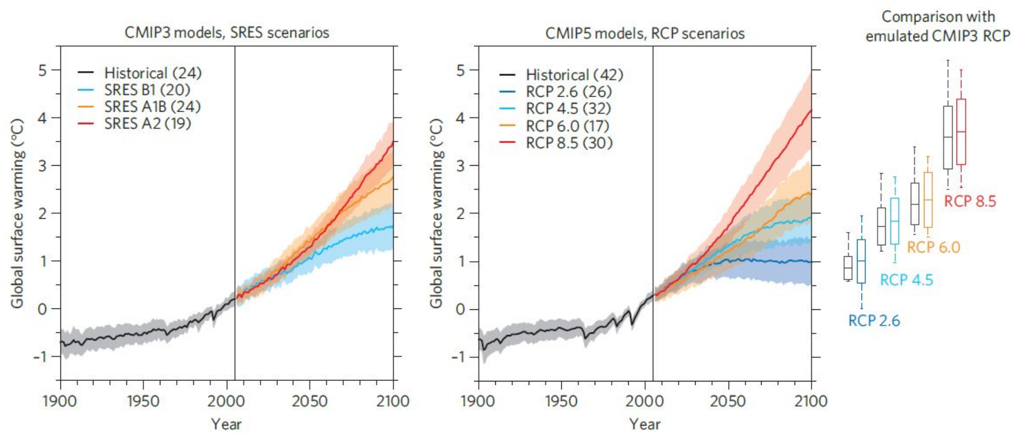

Figure 5. Global temperature change and uncertainty. Global temperature change (mean and one standard deviation as shading) relative to 1986-2005 for the SRES scenarios run by CMIP3 and the RCP scenarios run by CMIP5. The number of models is given in brackets. The box plots on the right are given for 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.

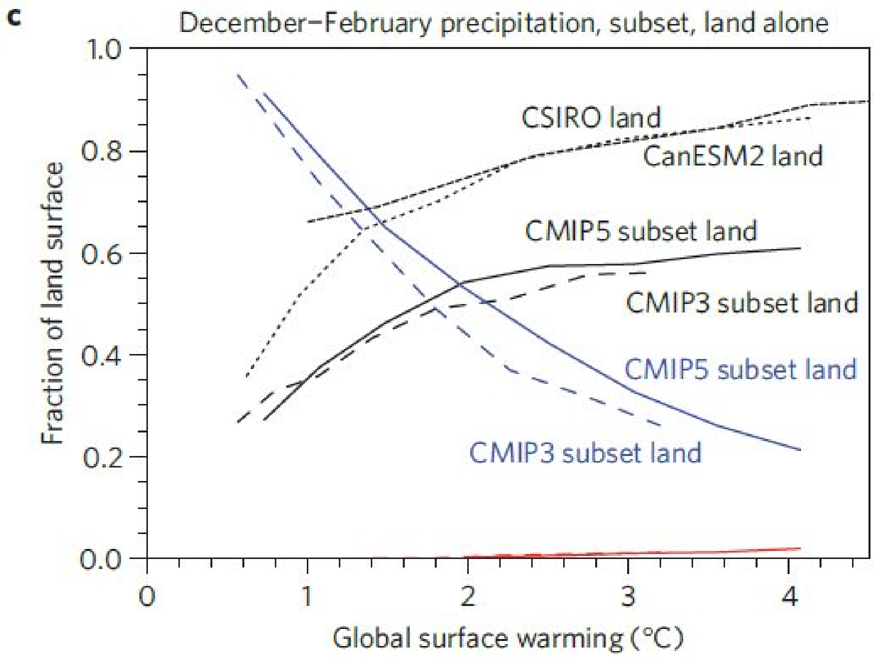

Figure 6. Model robustness for precipitation for land. Fraction of the Earth's land surface with high robustness (black, R>0.8), no significant change (blue) and inconsistent model responses (red) illustrated for December-February precipitation change for CMIP3 and CMIP5. Fractions are shown as a function of global temperature change rather than time, which largely eliminates the differences in scenarios. (Knutti and Sedlack, 2013) However, this graph is not straightforward or efficient in science communication.

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.

Moreover, although this visualization is scientifically rigorous and rich in information, but it does require a bit of effort to be able to interpret it correctly. Thus, we choose not to follow the style of this visualization to represent robustness, but to follow the model-spread type visualization as shown in Figure 5.

Objectives and Contribution

Our project uses novel data visualization techniques to communicate clearly and effectively the impacts of climate change on Great Lakes Tributaries, as well as the robustness of our model. Our objectives are to:

1. Improve upon the time series of model projections visualization style (Figure 2) to more effectively communicate climate change impacts to streamflow and temperature of Great Lakes tributaries. This is done through displays of model uncertainty that take into account color, data-ink ratio, and scale. We 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.
2. Explore improvements to the map-based climate change impacts visualization (Figure 3) to more effectively communicate climate change impacts to Great Lakes tributaries to the general public.
3. Improve upon the model reliability and tuning visualization style (Figure 4) to more effectively communicate the credibility and reliability of climate impact simulations of Great Lakes tributaries. This is done using an interactive map with effective color choices.
4. Improve upon the existing visualization techniques for robustness analysis and the concept of model robustness (Figures 5 and 6).

The contribution of our work is a great improvement in the communicability of climate science to the general public. Our interactive map can be embedded in a webpage and made broadly available for access. This will allow the public to be more aware of climate change impacts to Great Lakes Rivers and better able to craft policies and management plans regarding ecological, social, and economic uses of the changing rivers. The visualizations of model robustness and uncertainty will also help the public understand what the outputs mean, which could deter extreme reactions including outright denial or anxiety about apocalyptic events.

Data and Method

Data Sources

The datasets for these visualizations are 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 can be used until the final projections become available in 2020.

For the model robustness 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 used 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. The latitude, longitude, and evaluation (dr) data are numeric, while the calibration station data is nominal but stored as a number.

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 include monthly projections of streamflow, water temperature, and suspended sediment load for the Basin's rivers from 1950-2099 (e.g. Table 2). There is one set of projections for each climate model in the CMIP5 ensemble (n=20). The data for year and month are temporal, while the model output data are numeric.

Table 2. Head of time series streamflow data for the Maumee River (Station #174), displaying data from three of the 20 CMIP5 climate models.

Year	Month	access1-0	bcc-csm 1-1	canesm2
1950	1	20.69	30.51	5.464
1950	2	44.43	38.14	13.68
1950	3	51.99	14.11	84.27
1950	4	91.97	10.64	74.95
1950	5	44.1	34.95	81.64
1950	6	112.5	56.41	69.58
1950	7	139.7	42.68	19.45
1950	8	42.84	21.9	21.13
1950	9	16.18	32.95	13.61
1950	10	39.65	12.22	10.76

The data for the map of climate projections has been extracted from a large dataset as the percent change between the 2010s and 2090s decades, for one climate model (NORESM1-M). For example, a percent change in streamflow of 40% means that the model predicts average streamflow will increase by 40% between the 2010s and 2080s.

Table 3. Head of climate change projections data, showing the simulated % change between the 2010s and 2090s using the NORESM1-M climate model.

Station	Latitude	Longitude	Streamflow (%)	Temperature (%)	Sediment (%)
1	49.59799	-87.9642	20	54	149
2	48.92608	-87.6899	23	58	184
3	48.90472	-88.3759	31	55	158
4	48.84803	-86.607	19	54	71
5	48.82279	-88.5319	31	55	229
6	48.7782	-86.8839	21	55	171
7	48.77282	-86.2937	24	51	73
8	48.68789	-86.2129	21	52	61
9	48.53094	-89.5939	22	55	100
10	48.29189	-89.8084	16	59	80

Procedures

Our overall procedures follow these principles:

1. Accurate: accurately represent data. In other words, we try to use perceptually accurate graphical encodings.
2. Informative: keep low ink-data ratio, unless the extra ink will enhance experience or perception.
3. Straightforward: choose semantic meaning of colors, unless it interferes with the delivery of information or confuses our viewers.
4. Perceptual: choose perceptually uniform colormap unless it is not important.

More details about choices of methods and graphical encodings are case-dependent and will be explained in the “Results” section.

Results and Insights

Time series charts of model projections

The time series charts display the monthly streamflow and water temperature projections for the Maumee River (Station #174) from 1950-2100 (Figures 7.a., 7.b.). To make our model robust, we used 20 different options for climate models in these visualizations. Station # 174 has the most complete historical data and thus is a suitable proxy for model tuning. Prior time series visualizations had just displayed the data from separate models as different-colored lines (e.g. Figure 2). However, this uses a lot of unnecessary color ink and creates problems where the models overlap. We improved upon this by visualizing the outputs from the 20 models as translucent grey lines, displaying the ensemble mean as a thick dark line, and then creating a colored envelope of standard deviation around the mean. This quickly allows a viewer to see the typical projection as well as the uncertainty. A viewer will be able to see that these model projections converge nicely through the colored envelope and this is a piece of evidence for the robustness of the projections.

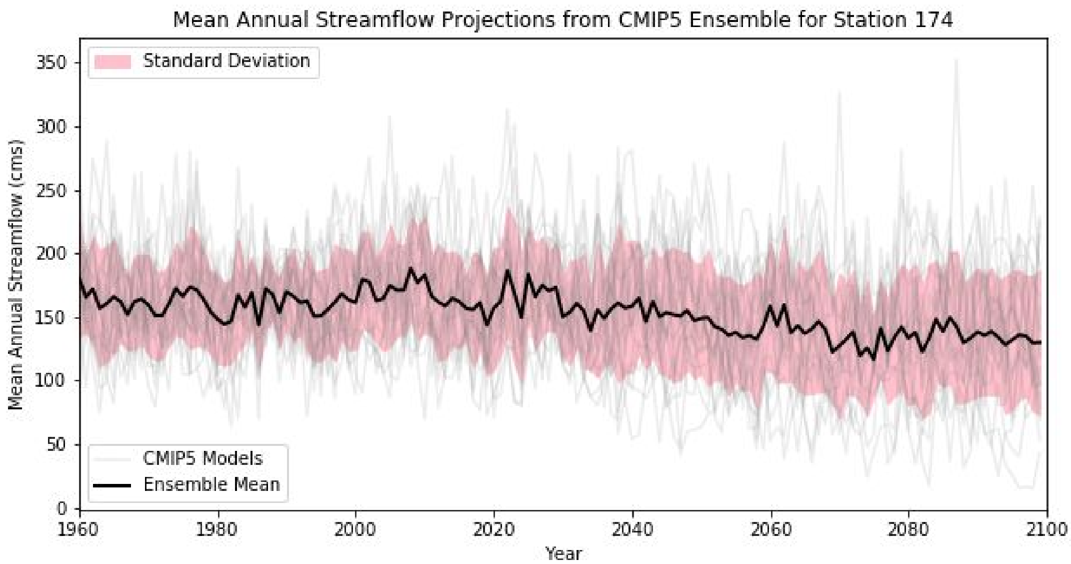

Figure 7.a. Time series charts of model projections: Mean Annual Streamflow Projections from CMIP5 Ensemble for Station 174.

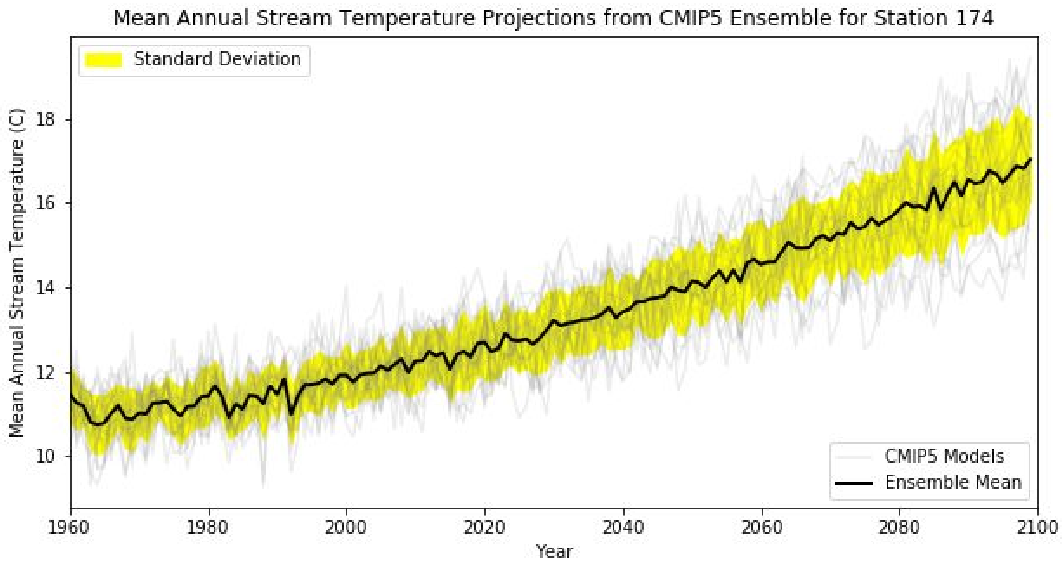

Figure 7.b. Time series charts of model projections: Mean Annual Stream Temperature Projections from CMIP5 Ensemble for Station 174.

We made the charts using python package MatPlotLib, with data stored as Pandas dataframes, and processed using Numpy. We experimented with different aspect ratios and X-Y limits for the figures, and decided on the current style because it fits neatly on a page and does not lead to misguided interpretations of overly steep or flat slopes. This way, the visualizations are informative to the general public while not being manipulative. The colors of the envelopes have to be of low saturation, so that they do not obscure the mean or completely block the individual model projections.

Additionally, we used the same data to make a time series chart, but with box plots to show decadal summaries (Figures 7.c. and 7.d.). This allows for a quick and easy interpretation of trends and variation in the data.

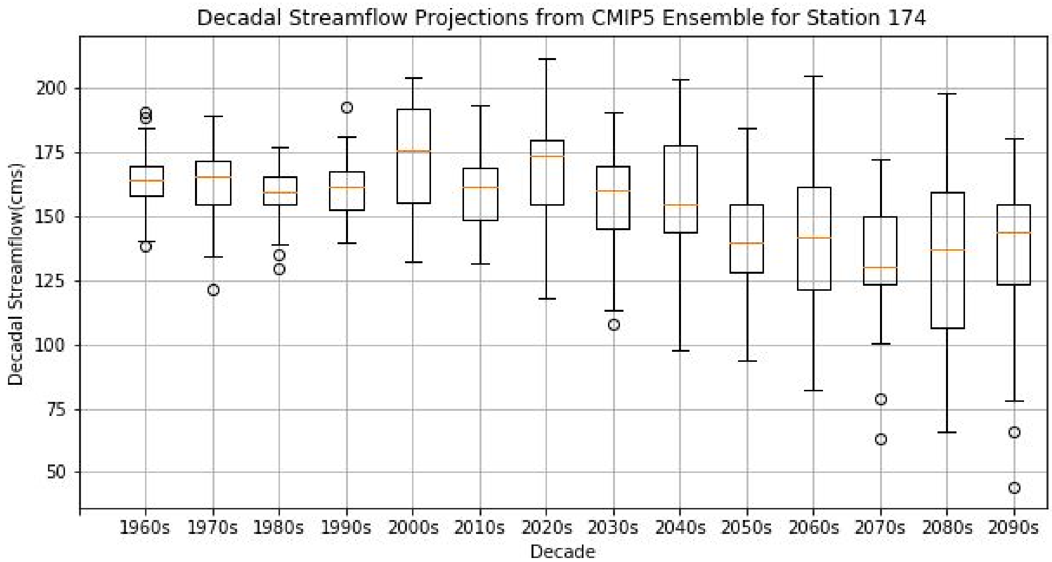

Figure 7.c. Time series charts of model projections: decadal streamflow projections from CMIP5 Ensemble for Station 174 with boxplots.

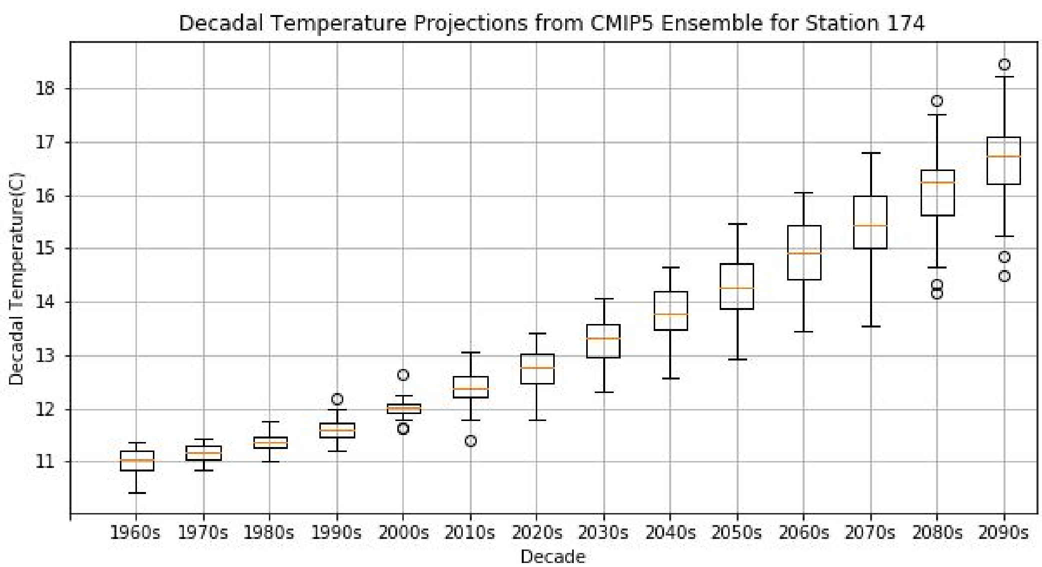

Figure 7.d. Time series charts of model projections: decadal stream temperature projections from CMIP5 Ensemble for Station 174 with boxplots.

First of all, we turned the grid on so that viewers can compare the boxplots with x-axis and y-axis fairly easily. Second, the boxplots also allows viewers to see outliers in each decadal period. Thirdly, these boxplots are different from the Figure 7.a and b in the following aspects: (a) decadal, instead of yearly, dispersion is captured; (b) boxplots trace the median and quartiles of distributions, rather than the mean and standard deviations. This is why we think time-series charts with boxplots are necessary and complementary to the previous charts. These were also created with Python package MatPlotLib.

Data map of climate change impacts

The filled contour maps of climate change impacts were created using the tricontourf() function of MatPlotLib. We originally chose the “seismic” colormap to visualize the projected streamflow changes because it is divergent, with red colors representing an increase and blue colors representing a decrease (Figure 8.a.). We normalized the colormap so that the zero is white and added a color bar for reference. Although the use of different shades of colors could cause some confusion due to the effects of adjacent shades on interpretation (e.g. darker adjacent colors make the data appear lighter), we deemed this an appropriate choice because of the ease of quickly visualizing patterns. For example, in Figure 8.a, viewers can easily see that the north is becoming wetter while the south drier. Then we realized that we need to put some reference to this chart because the longitude and latitude information is not enough for viewers to immediately understand where this region is. The best way to do it is to plot natural landmarks such as the lake shore and rivers, and the state and country boundaries.

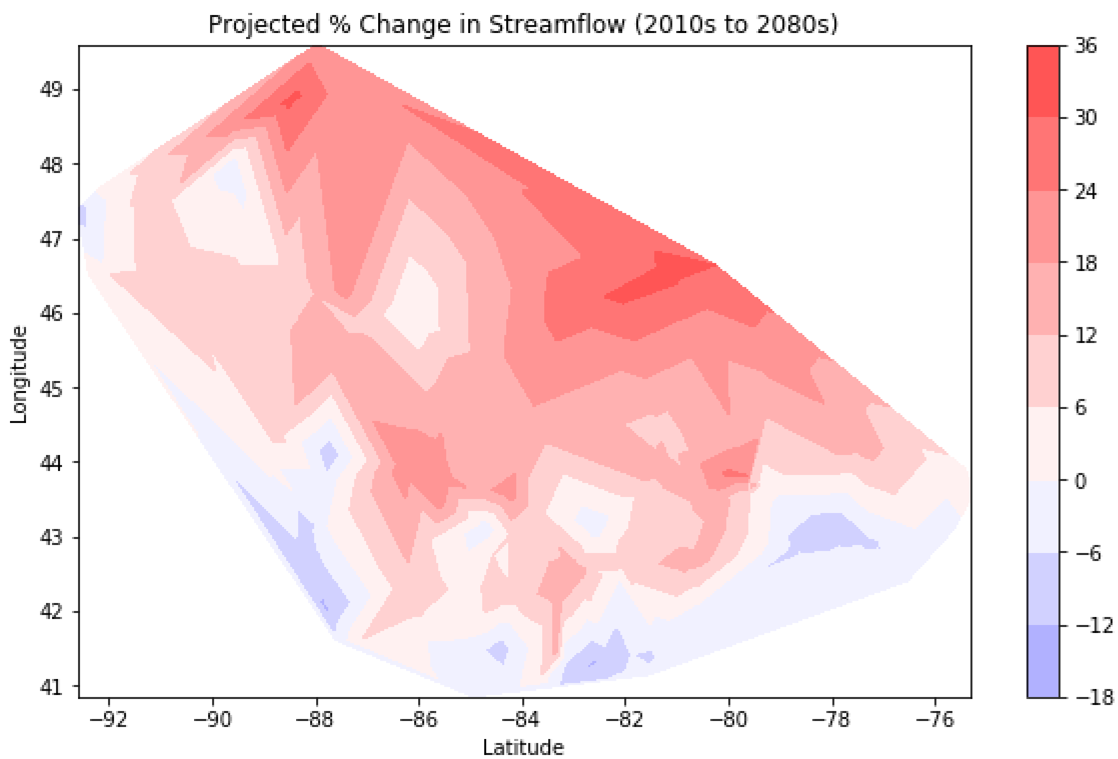

Figure 8.a. Data map of climate change impacts: original contour-fill of projected streamflow change with “seismic” colormap.

Based on the filled contour map, we plot a Basemap as background. With the outlines of lakes and rivers in the Great Basin area, viewers will be able to relate the projected changes to the actual geographical location. We think this map background enhances the delivery of information. For anyone who is familiar with the Great Lake region, it is fairly easy to recognize this plot. Also we change the colormap to “coolwarm” and adopt the semantic meaning of red and blue (Figure 8.b.). That is to say, in the data map of projected stream flow, red means the area is getting drier while blue means the area is getting wetter. In the projected temperature, red simply means the rising trend of projected temperature (Figure 8.c.).

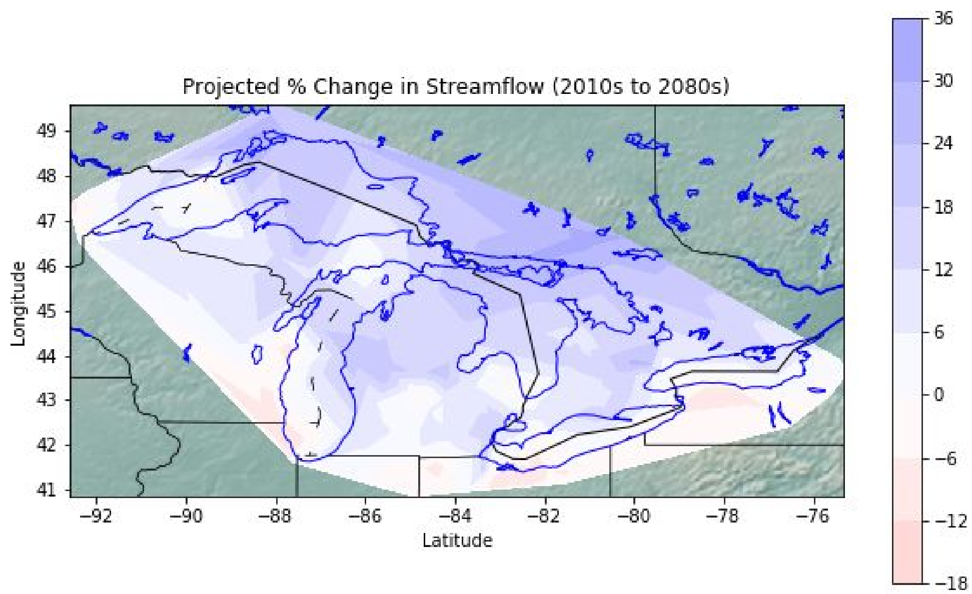

Figure 8.b. Data map of climate change impacts: contour-fill of projected streamflow change with “coolwarm” colormap based on real map. It is shown that the north is getting wetter than the south.

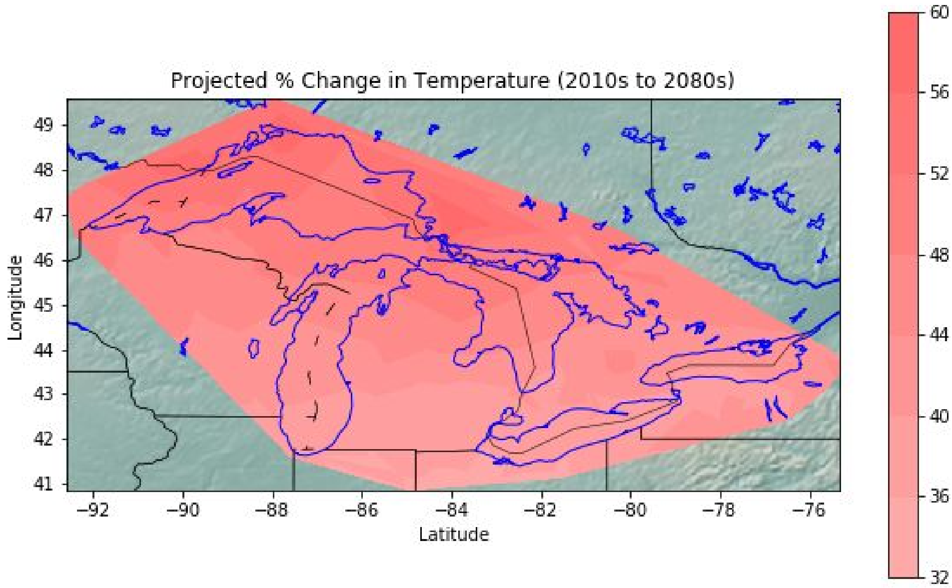

Figure 8.c. Data map of climate change impacts: contour-fill of projected temperature change with “coolwarm” colormap based on real map. It is clear that the temperature is rising overall in the Great Lake Basin area.

Figure 8.c still uses “coolwarm” colormap because we want to keep the consistency between Figure 8.b and 8.c. We flipped the color map in order to keep the semantic meaning of color, i.e., red meaning warming. Admittedly, the pattern in Figure 8.c is not as distinguishable as the one in Figure 8.b, but this is mainly a result of our intention to keep the colormap consistent. Additionally, instead of using percent changes, we could use the changes in absolute temperature values, which might be more straightforward and intuitive to some viewers.

Interactive map of reliability and model tuning

We began our project by making the map of model reliability and tuning. We used the package Pandas to store data, Numpy for numerical processing, and MatPlotLib to create plots. We used points to display data because each calibration station with data exists at specific latitude and longitude coordinates.

Originally, we used the Altair package to explore visualizations of the model reliability data (Figure 9). However, this package was limited in its interactivity with basemaps, callouts, and widgets that would be helpful for the public to understand the findings. Thus, we used the packages ipyleaflet and ipywidgets to make the map interactive. To do this, we needed the package ColorMap to convert the format of the colors from RGB to Hex digits. Finally, our data is presented in the Mercator projection, which uses straight and orthogonal latitude and longitude. The interactivity and sequential colormap are both major improvements over the existing model reliability visualizations (e.g. Figure 4).

Figure 9. Exploratory visualization of Great Lakes Basin station data in in Altair map.

We originally chose the “viridis” colormap because it is sequentially uniform, so thus a perceived difference in color corresponds to a similar actual distance between values (Figure 10). We eventually decided on “autumn” as the color map for the dots representing the water station, although we understand that “autumn” is not as perceptually uniform as “viridis” (Figure 11.). The interactivity of the map allows the viewers to check the numerical information of a specific water station and thus the perceptual uniformity does not seem to be very important. Second, viridis colormap, as a colormap ranging between blue-green-yellow, is not very aesthetically attractive. Given these two considerations, we eventually decided on “autumn”.

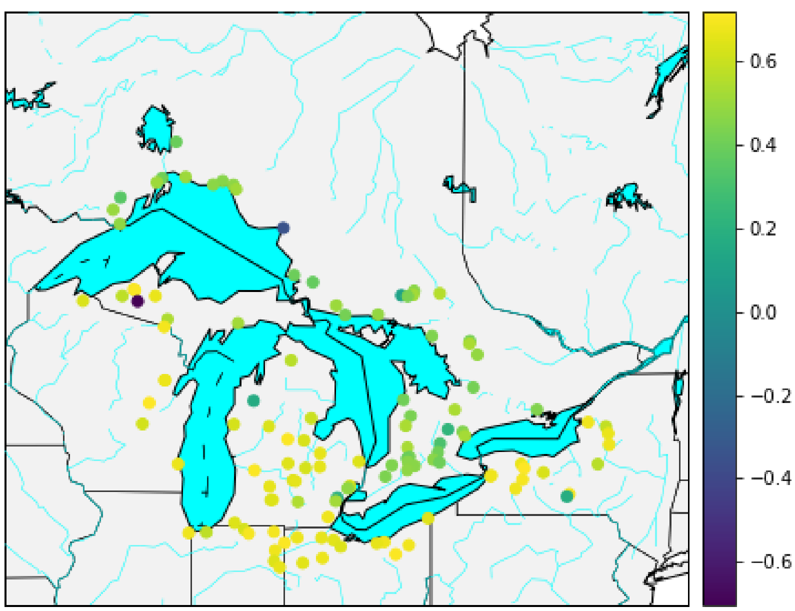

Figure 10. Exploratory data visualization of Great Lakes Basin stations using the Basemap package and perceptually uniform “viridis” colormap.

There are a few features of interactivity. First, this map supports drag and zoom, which comes with the package. Second, this map supports pop-ups. Once the viewer clicks on one dot that represents a water station, an information box will pop up, showing the water station number, its coordinates, and some of its other associated data. This map also comes with a widget control. The viewer can access the information box by typing in the water station number in the dialogue box in the corner. For the scope of this project, we did not proceed to learn how to host this map on a website. However, we have a video demo hosted on YouTube.com.

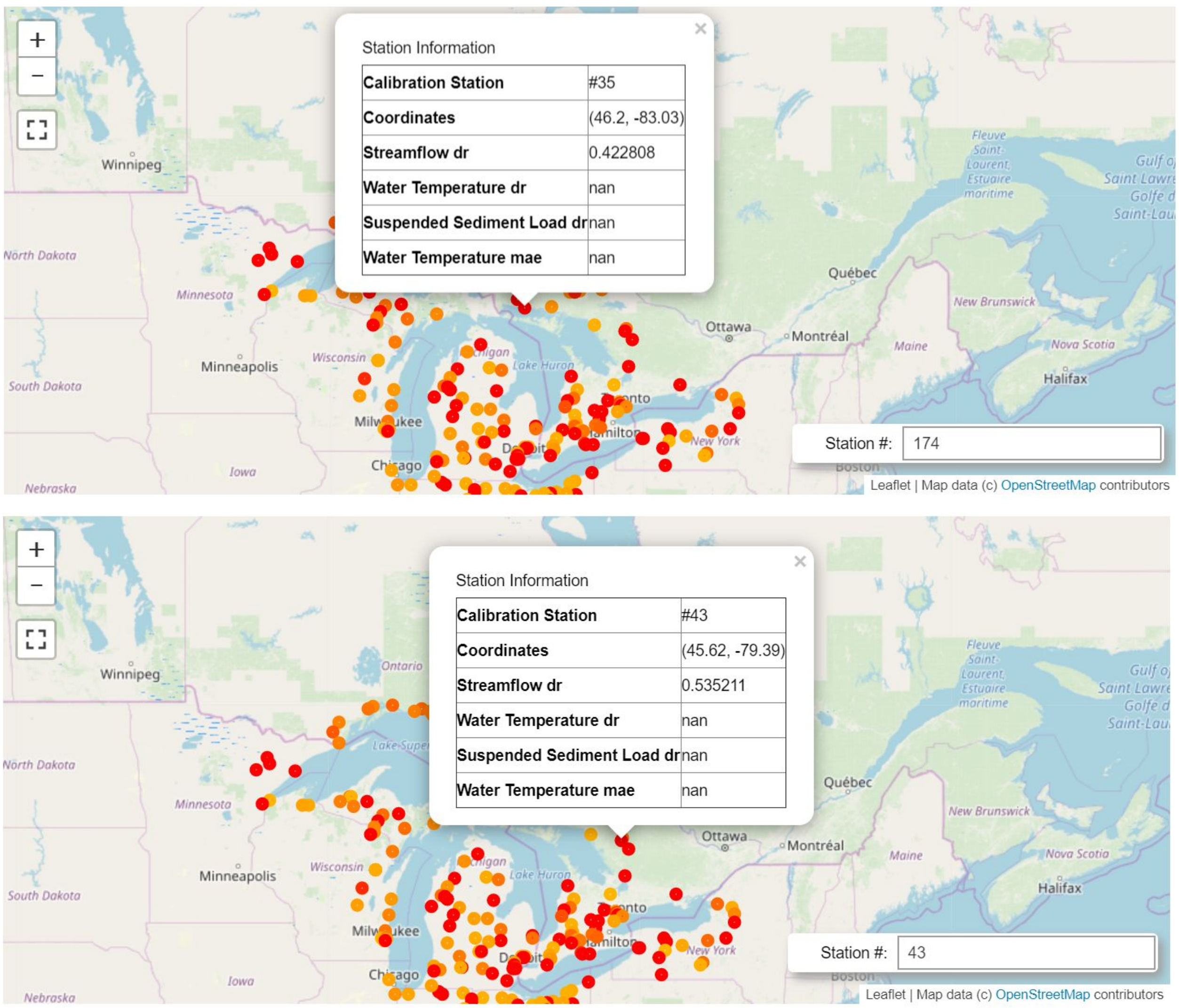

Figure 11. Interactive map of reliability and model tuning: (Screenshots) This interactive map supports 1) dragging and zooming, 2) clicking on a station to see the information, 3) widget control in the corner to find a particular station. Here's the demo for this map: https://youtu.be/Y_OmiaRec5E.

Discussion

Through project, we have learned more about our data set and also improve our skills in effectively visualizing these results. First of, from these visualizations, we learned that the models were robust and performed well because their results largely converge on a similar trend. Of course, this could be done by merely calculating the standard deviation of projections or model spread, but the numbers are not as straightforward as a visualizations. Second, we also learned that water temperature will be a major concern into the future, as projections for the Maumee River have strong agreement among models for a temperature increase of several degrees Celsius. This could be problematic for fish in the river (Ficke et al. 2007). This information is also best represented through charts and graphs. We also produced convincing evidence via the interactive map, so that viewers can check and verify the statistics of each water station. The visualizations also allow the public to better understand the impacts of climate change. This gives decision makers an improved ability to manage the rivers to adapt to climate change risk, improving the surrounding ecologies, communities, and economies into the future.

We have also familiarized ourselves with some Python packages, such as Pandas, Numpy, MatPlotLib, and Ipyleaflet. These packages are of general purposes. Our experience with this project can be easily transplanted to other data set that involves maps or time series, or multi-model projections.

Conclusion

Effective visualizations of climate change impacts and model robustness are essential for communicating the science to the general public In this work, we used the datasets from the Indiana University research project “Hydrologic Modeling for Sustainable Management of Great Lakes Basin Rivers Impacted by Climate Change,” funded by the IU Sustainability Research Development Grant, to produce a series of visualizations that aid in communication of the research. These included 1) time-series visualizations of projected temperature and streamflow changes, showing model robustness and uncertainty, 2) A data map of climate change impacts across the Basin, and (3) interactive data map showing the information of all the water stations around the Great Lakes Basin areas and model tuning. These visualizations allow the climate science to be communicated clearly and concisely.

Future Work

As to present project, a few further improvements can be considered: (1) the prototype of the interactive map (Figure 11) should be tested, published online, and made public to the general public. (2) A non-interactive 2D visualization for model tuning and dr-statistics should be considered for the purpose of scientific publication.

Further, climate models are continuously being improved. As we write, a new ensemble of climate models known as the CMIP6 group have been created for a coarse global scale. Future research should downscale these models using historic weather data to be more applicable to regional or local hydrology studies. Future work should involve an update to our visualizations once these newer data become available. Additionally, the best management practices for climate science data visualizations should be shared with researchers around the globe, to help make visualizations that are more understandable by the general public.

Acknowledgements

Indiana University Geography Department
Indiana University Sustainability Research Development Grant
William R. Black Fellowship
National Science Foundation [grant numbers DBI-1564896 and CNS-0521433]
Indiana University Pervasive Technology Institute, Lilly Endowment, Inc., Indiana METACyt Initiative, and Shared University Research Grants from IBM, Inc. to Indiana University.

Appendix

Jupyter Notebook Script:

You can access the data and script for this project here:
https://drive.google.com/drive/folders/1AHjLFB-BoDLwk47-dX-GUMoAiqFUxKqo?usp=sharing

References

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.