Visualizing ENSO

Logan Jones, Sina Soltanieh, Alexander Schad

Northeastern University

1. Abstract

The El Niño Southern-Oscillation (ENSO) has long been known for its strong influence on global weather. Although localized to the Pacific Ocean, its effects can be seen across the globe, with notable effects around the Mississippi River Basin (MRB). In 2017, Muñoz and Dee found that the warm phase of ENSO preconditions the MRB to flood vulnerability [1]. Given the economic costs associated with natural disasters in this heavily populated region, understanding the connection between ENSO and flooding is of paramount importance—allowing easy visualization and exploration of this connection is the focus of this paper and the resulting website that was created.

Keywords: ENSO, climate change, geospatial visualization.

Github Repository

2. Introduction

The MRB is one of the most heavily engineered and navigated river valleys in the world. It is a key driver of the American Midwest economic engine, and its historical susceptibility to flooding has been known for its costliness.

The El Niño Southern-Oscillation (ENSO), although reasonably well-understood by climatologists and meteorologists, continues to show a surprisingly notable impact on regions across the globe. In 2017, its connection to the lower MRB was made clearer than ever after the publication of, "El Niño increases the risk of lower Mississippi River flooding," by Muñoz and Dee [1]. This study showed that the warm phase of ENSO showed a clear, causal impact on the susceptibility of the lower Mississippi to flooding. More specifically, in the 6-12 months preceding a given flood within the region, the authors found that El Niño's addition of abnormally high rain levels in the region reduced the capacity for soil infiltration by water. At the time a flood occurs, there is little room for excess water capacity from extreme rains to drain, leading to devastating flooding throughout the lower Mississippi.

Although this paper was crucial in illustrating the connection between ENSO and lower MRB flooding, the public understanding of the connection and political responses to it have been limited. In order to make its findings more tangible and accessible, we worked to develop an interactive web-based visualization that allows users (citizens, policymakers, and scientists) to geospatially explore the complex connections between El Niño and the lower Mississippi. By making this phenomenon more accessible to the public, we hope to highlight the critical importance of building a more resilient Mississippi infrastructure that is prepared for the growing impact of climate change. The visualization website that we created is designed to allow easy exploration of ENSO and MRB connections and subsequent insights about the important role that ENSO plays in influencing MRB weather conditions.

Webpage

Demo Video

Project Slides

3. Related Work

1) El Niño increases the risk of lower MS River flooding [1]
This paper describes the climatic relationship between ENSO and lower MRB flooding, which we hope to better visualize.

2) Climate Reanalyzer [2]
This website contains many geospatial visualizations related to climatology and weather that will inspire our work.

3) Real-time 3-D El Niño/La Niña visualizations [3]
This IEEE paper tackles one aspect of the project that we hope to address – how to visualize the ENSO spectrum.

4) El Niño [4]
This is a NASA comprehensive guide on the El Niño phenomenon. It will assist in better understanding ENSO science.

5) Interactive web access to real-time El Nino/La Nina data [5]
This explores a tool made to interactively visualize El Niño and will be a useful reference.

6) Importance-Driven Time-Varying Data Visualization [6]
This IEEE paper suggests different visualization techniques for different temporal trends, a useful concept for our work.

7) Correlation Study of Time-Varying Climate Data [7]
This IEEE study also focused on temporal curves as effective methods to explore correlations in climate datasets.

8) Visualization of Geospatial Time Series [8]
This EuroVis research introduces an approach for visualizing large geospatial time series, an important concept for our research.

9) Geospatial information visualization [9]
This is a chapter from Meeks's book, D3.js in Action. It provides detailed information on how to visualize geospatial data using D3.

10) D3.js: Introduction to Mapping [10]
This is a guide on using D3 to visualize geospatial data. Although it works with JSON, it will be useful for our purposes.

4. Partner

Our partner for this project is Dr. Samuel Muñoz, assistant professor at the Northeastern College of Science. We are collaborating with Dr Muñoz in order to augment the work of his lab in better understanding the connection between ENSO and Mississippi River flooding and to improve the accessibility and visibility of his group's research. Dr. Muñoz and his lab plan to use our website to promote their research on their website and at conferences.

In our conversations with Dr. Muñoz, he has been an excellent resource in learning more about the science behind ENSO, and we had fruitful discussions about how to visualize the phenomenon and its connection to the Mississippi river basin.

We were surprised by the wide availability of data related to sea surface temperatures and soil moisture. Although we had already explored and found data applicable to our project, Dr. Muñoz was able to show us a lot more data options and sources that we could potentially use.

Dr. Muñoz also effectively communicated what exactly his lab was looking for in a visualization, specifically discussing the types of interaction, data, and aesthetics for which he was looking. This made the process of creating a final product much easier. After a few feedback sessions and iterations, we were able to create a product that was in line with the lab's needs.

5. Data

Data used for the El Niño project is sourced from the National Oceanic and Atmospheric Administration (NOAA) and the United States Geological Survey (USGS) and is publically available. File format of the data is NetCDF, however, Dr. Muñoz has an R script available to flatten the file into a more readable format.

Additionally, Dr. Muñoz has provided us with several data sources that could potentially be utilized for our project: historical sea surface temperatures [11], historical precipitation [12], historic soil moisture [13], and Mississippi river levels [14].

6. Task Analysis

Task ID	Domain Task	Analytic Task (low-level, "query")	Search Task (mid-level)	Analyze Task (high-level)
1	Explore connection between ENSO and MRB conditions	Summarize	Explore	Discover
2	View Pacific SST during ENSO cycle	Identify	Lookup	Discover
3	Compare MRB precipitation levels across ENSO spectrum	Compare	Browse	Discover
4	Compare MRB soil moisture levels across ENSO spectrum	Compare	Browse	Discover
5	Compare MRB river levels across ENSO spectrum	Compare	Browse	Discover
6	Examine environmental conditions over the ENSO time cycle	Compare	Browse	Discover
7	Examine ONI index across ENSO time cycle	Identify	Browse	Discover

Our visualization will be developed primarily for discovery about ENSO's several indicating factors to show flooding in the Mississippi River Basin. The visualization is motivated by existing theory on ENSO, and will be used for users to verify whether conjectures about the El Niño and La Niña phenomena are true (discovery presentation of the known). It will not be telling a direct story, but rather allowing the user to move around to discover and verify hypotheses. Thus, "discover" or exploratory visualization will be our main type of consumption.

The primary user of visualizations would be scientists and those working in Dr. Muñoz's lab. Also, scientists and researchers outside of the lab would use these visualizations to help support their findings.

7. Execution & Design Process

Plan of execution follows from our goal of creating an effective and immersive visualization to display the effects of El Niño and La Niña. We will be creating the visualization using Data-Driven Documents (D3.js), which is a JavaScript library that binds data to the DOM and applies transformations and events using HTML, CSS, and JS.

Considering most of the data is geospatial, we utilized our task analysis and task abstraction in order to effectively display this data for our target user. Utilizing libraries D3 provides to transform geospatial data (GeoJSON, TopoJSON) may be helpful as well as using functions to create Mercator, Mollweide, orthographic, and satellite projections in D3 [11].

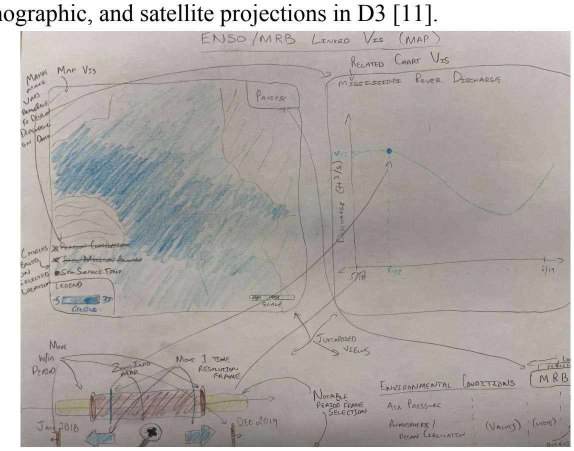


Figure 1: First sketch featuring maps

In terms of visual encodings, our visualization will include four juxtaposed views. On the right side, we will have two geospatial maps. The top will be a map for a selected variable (controlled via a radio form on the left) such as precipitation or soil moisture of the Mississippi River Basin and the bottom will be a map of the Pacific with sea surface temperature values. Both of these maps will use circles as marks, each representing a latitude longitude pair on the map, which will implement details-on-demand via hovering. We will then look to use a library to implement a color measure based either on a distance-based gradient (continuous saturation channel) or different color zones grouped by value (ordinal hue channel). After discussing the desired views with our project partner, we have satisfied all requirements with these views.

For the left side, we will have a timeline for our time-series time axis. On top of the timeline, there will be a line graph for the specified variable and below the timeline, the ONI index line graph will help to distinguish ENSO events. The line charts use line marks with vertical and horizontal position channels. The horizontal axis will be time for both charts, while the vertical dimension is the selected variable (top view) and the ONI index (bottom view). The ONI index chart additionally contains an area mark with a categorical hue color channel from the line to the horizontal axis which indicates whether the event is El Niño (red) or La Niña (blue).

For our interactions and interactive components, all four views will be connected via the time-series timeline axis. Moving the pointer along this axis will move a point along the line charts which will indicate the corresponding variable and ONI index values. It will also shift the time of the two geospatial maps to display new values for the shifted time.

After reviewing our task analysis, all four views and the time interaction between them described above are necessary to help users explore the relationship between ENSO cycles and MRB environmental conditions. Some nice to haves would be a zoom or move feature on the map and an advanced slider with start and end time periods that can be manually set or selected from a list of known ENSO periods.

Implementation for necessities will be done using the D3 library. Interactions between views will be handled using d3-dispatch. The challenge to create a color map for geospatial marks due to latitude and longitude resolution is currently being looked into and we plan on trying a voronoi tessellation provided by the d3-delaunay library (as suggested by Prof. Cody Dunne).

Our visualization does not require any server side components.

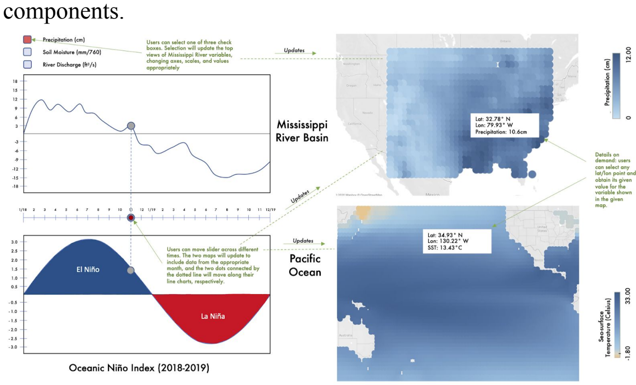


Figure 2: Final Digital Sketch

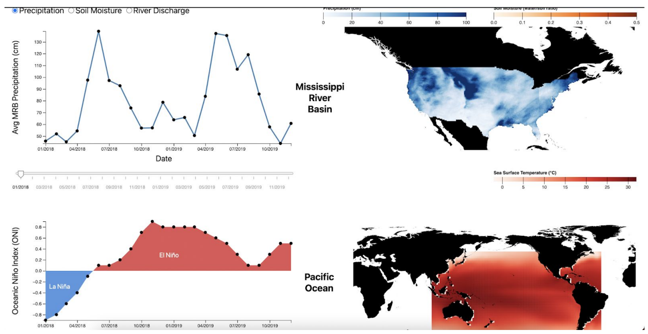
This final digital sketch was a good background to start developing our visualizations but through iteration and trial and error, we eventually tweaked our original design to create visualizations that were more user friendly and effective. Originally, the points on the maps were circles, but we found that this did not translate well onto either map as there was white space in between points. To fix this we changed the points into square areas to minimize white space in between points.

After receiving feedback from usability testing, we were able to implement some improvements to our original design. We learned that the zoom feature which we originally intended for users to be able to zoom in farther into the maps to get more accurate data ended up being a nuisance and not necessary. To fix this we bounded the maps so users can't pan or zoom but instead could pay attention to the area that is most important which is the Pacific Ocean for the Sea Surface Temperature map and the Mississippi River Basin for the top left map. We also made it more clear to users how the visualizations were connected to the slider by having the slider and the two line graphs have the same x axis dates. Previously, the slider was numbered 1 to 24 which was confusing to some users who didn't understand what the slider was supposed to represent. Streamlining the x axis and slider enables users to better understand the connection between the graphs and the user interactive slider.

Other than these changes our final product ended up being fairly close to our final sketch. Our final visualization enables users to be able to fulfill all of the tasks defined above to our partner's needs. The four quadrants enable users to be able to see all of the information on their screen without scrolling, or zooming. Our choice of using the position channel in two views and color in the other two are the best possible choices since position the most accurate channel cannot be used for maps as it is inherent to its structure. We use points to showcase specific data at a specific point in time or place, and use the line mark to see the change in time of the line graphs. To get more accurate details on demand, users are able to click points in all four graphs and can interact with the slider to see the change in all visualizations over time. Overall, we believe our design choices and execution were able to satisfy all of our goals.

8. Visualization Design

In terms of the appearance of our visualizations, we haven't deviated from the final design sketch too heavily. We have four visualizations each with their own quadrant with a slider in between the ONI chart and the line graph on the left. The top left features the line graph, the bottom left is the ONI chart, the top right is the MRB map while the bottom left is the SST heatmap. The colors red and blue are used to contrast El Niño and La Niña in the bottom left ONI chart. Also, we chose red to represent El Niño to represent warm waters while La Niña represents cooler waters so it is blue. We purposefully chose colors that were not too heavily saturated in order for the graph to be easier on the eyes. We also changed the color for the heatmaps so that each would stand out by being distinct colors. We chose a brighter orange to represent soil, red to represent the temperature of the water, and blue for the precipitation. In terms of interaction, the main way users interacted with this visualization, other than details on demand for points in all four views, was through the time-series controller slider bar. Each tick on the slider bar was discrete and represented a month between 2018 and 2019. Moving and releasing the slider bar updates the monthly averages for the points on the map views to the selected month. Note that due to the sheer size of this data (about 35 megabytes), we were unable to smoothly update the points as the user dragged the slider, due to high render times which would lag the visualization, so we settled for updating when the user released the slider button. For the line chart displaying environmental data, radio buttons make it easy for users to switch between soil moisture, precipitation, and discharge data, which updates both the line chart and the Mississippi River Basin map view to display point averages for the selected measure (discharge data is not geospatial, so a message is displayed to the user saying so). We ended up changing the decision to zoom and pan, as data points were based on specific latitude longitudes and not boundaries, there wasn't any additional info being discovered when zooming in, and it was easy for users to get lost in the zooming and panning due to lack of showing extra map detail, which would result in the user having to refresh the page. Thus, we decided it would be best to remove this interactive feature.



Screenshot of current state of visualization.

9. Discussion

The primary finding of our study was that an interactive visualization is far superior to a static visualization because users are able to explore the data and find insights on their own. In the case of ENSO and MRB, this is especially important—although scientists have already found a clear connection between the two that justifies targeted preparation to prevent flooding, policymakers have done little to take action. This project makes the connection between ENSO and MRB far more salient, acting as an important means of scientific communication in an effort to spur policy action.

This project was also an important lesson in dealing with the inevitable roadblocks that come up in any complex project. Despite thorough planning, various technical issues related to creating or formatting visualizations in D3 arose throughout our project. Utilizing our resources—the internet, TAs, the professor—was crucial to overcoming these hurdles and a reminder of the importance of communication in collaboration.

Finally, this project highlighted an important limitation inherent to most interactive geospatial visualization projects—speed. Because of the sheer size of geospatial data, specifically projections and lat/lon points, geospatial projects must be able to consume and handle large amounts of data. Keeping our interactions fast was a major hurdle that we feel like we reasonably dealt with, but the website's speed of updating is not ideal. This is a limitation inherent to geospatial project, but in the future, we hope to explore algorithms that might increase speed in order to improve usability.

10. Conclusion

In this project, we were able to successfully create a set of interactive, linked visualizations that allow users to better understand the connections between ENSO and MRB weather conditions. This work will be useful as the Earth Surface Systems Lab seeks to spread the findings of their research online and in conferences. In the future, our group will continue to refine the visualizations with Dr. Muñoz and his team—improving update speeds, aesthetics, and data included in the dashboard. Ultimately, we hope that this research will contribute to the scientific understanding of citizens and policymakers interested in the importance of ENSO on global weather conditions.

11. Acknowledgements

We would like to thank Dr. Muñoz and the Earth Surface Systems lab for the collaborative efforts in explaining ENSO, pointing us to data, and working to design a visualization that suits their needs.

12. References

1. Samuel Muñoz and Sylvia Dee.. El Niño increases the risk of lower Mississippi River flooding. *Sci Rep* 7, 1772 (2017). doi: 10.1038/s41598-017-01919-6
2. Climate Reanalyzer (https://ClimateReanalyzer.org), Climate Change Institute, University of Maine, USA.
3. W.H. Zhu, E.F. Burger, D.C. McClurg, D.W. Denbo, N.N. Soreide, "Interactive web access to real-time El Nino/La Nina data from the TAO buoy network in the tropical Pacific", *OCEANS '99 MTS/IEEE. Riding the Crest into the 21st Century*, vol. 1, pp. 466-469 vol.1, 1999. doi: 10.1109/OCEANS.1999.805001
4. M. Carlowicz, S. Schollaert Uz, and J. Stevens. "El Niño," Nasa Earth Observatory, 2017. https://earthobservatory.nasa.gov/features/ElNino
5. McClurg, D.C. & Moore, Christopher & Soreide, N.N.. (1999). Real-time 3-D El Nino/La Nina visualizations and animations from the TAO buoy network in the Tropical Pacific. 2. 951 - 954 vol. 2. doi: 10.1109/OCEANS.1999.805001
6. C. Wang, H. Yu and K. Ma, "Importance-Driven Time-Varying Data Visualization," in IEEE Transactions on Visualization and Computer Graphics, vol. 14, no. 6, pp. 1547-1554, Nov.-Dec. 2008, doi: 10.1109/TVCG.2008.140
7. J. Sukharev, C. Wang, K. Ma and A. T. Wittenberg, "Correlation study of time-varying multivariate climate data sets," 2009 IEEE Pacific Visualization Symposium, Beijing, 2009, pp. 161-168, doi: 10.1109/PACIFICVIS.2009.4906852
8. Köthur, P., Sips, M., Kuhlmann, J., & Dransch, D. (2012). Visualization of Geospatial Time Series from Environmental Modeling Output. In EuroVis - Short Papers. The Eurographics Association. doi: 10.2312/PE/EuroVisShort/EuroVisShort2012/115-119
9. Meeks, E. (n.d.). Chapter 7. Geospatial information visualization · D3.js in Action. Retrieved October 05, 2020, from https://livebook.manning.com/book/d3-js-in-action/chapter-7/
10. Hanson, Bailey & Seeger, Christopher. (2016). D3.js: Introduction to Mapping. https://www.researchgate.net/publication/310425269_D3js_Introduction_to_Mapping
11. NOAA. "Extended Reconstructed Sea Surface Temperature (ERSST) v5". Retrieved October 20, 2020, from https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v5
12. NOAA. "GPCC Global Precipitation Climatology Centre." Retrieved October 20, 2020, from https://psl.noaa.gov/data/gridded/data.gpcc.html
13. Physical Sciences Laboratory, NOAA. "CPC Soil Moisture." *NOAA Climate Prediction Center (CPC)*. Retrieved 26 October 2020 from https://psl.noaa.gov/data/gridded/data.cpcsoil.html.
14. U.S. Geological Survey, U.S. Department of the Interior. "Surface Water Data for USA: USGS Surface-Water Monthly Statistics." *USGS Surface Water Data for USA:*

USGS Surface-Water Monthly Statistics, 26 Oct. 2020, waterdata.usgs.gov/nwis/monthly/?referred_module=sw&site_no=07289000&por_07289000_79165=1872479,00060,79165,2008-01,2020-01&start_dt=2008-01&end_dt=2020-01&format=html_table&date_format=YYYY-MM-DD&rdb_compression=file&submitted_form=parameter_selection_list

15. Giusti, Michela, "How to convert NetCDF to CSV". Retrieved October 20, 2020, from https://confluence.ecmwf.int/display/CKB/How+to+convert+NetCDF+to+CSV

13. Group Charter

Group Purpose:

This group has formed in order to provide the Northeastern University Marine and Environmental Sciences Lab with accessible and interactive visualizations and to meet Dr. Samuel Muñoz's needs and goals. Our primary goal is to make the data and research learned from Dr. Muñoz's lab more accessible and friendly to promote scientific research and discovery.

Group Goals:

This group's main goal is to assist our project partner in visualizing the connection between ENSO and MRB conditions, with the target of assisting their larger research in its accessibility. In addition, we hope to create a project that we will collectively show off to future employers in order to showcase our skills and abilities in creating effective data visualizations for a partner/client. We are all willing to commit ourselves fully in order to get an A on this project and in this course.

Member Roles and Responsibilities:

Logan Jones - Group Leader and Communications Director

- Primary point of contact between Dr. Muñoz's lab and the group as well as primary communicator with the Professor and the TAs of the course.

Sina Soltanieh - Documentation Coordinator

- Confirms all of the project's documentation is up to date and in a proper format.
- Manages and organizes project data.

Alexander Schad - Meeting Facilitator

- Creates weekly meetings and facilitates group discussion on what needs to be done.

Ground Rules:

In order to effectively progress the project at a reasonable and timely pace, we agree to meet at least once a week at 11 AM every Thursday for 30 minutes while also allowing any ad hoc meetings to take place if we are available. During these meetings we will be on Zoom and will be able to communicate details on the project and any issues, deadlines that might be of importance. In order to make decisions and to handle dissenting views among team members, we will go by a majority rule vote on the matter. To hold each other accountable, we will bring up any issues or concerns during our weekly meetings and try to resolve them there. We expect an equal amount of participation and dedication to this project from each other in order for all of us to be able to complete project tasks in a timely manner.

Potential barriers and Coping Strategies:

Because all of our schedules are different, there could be some days or weeks where our workloads outside of this class are different. This means some group members might be able to contribute more than others on any given week. In order to alleviate any issues in scheduling or any imbalance in work put into the project between members, we will meet at the bare minimum of once a week in order to communicate any issues or concerns we may have with the project.

Week 5 Group Charter Update

By week 5, all group members have contributed to the project deliverables adequately. We have had no troubles with completing work on time for the deadlines. When creating the charter originally, we agreed to meet every Thursday at 11 am for 30 minutes, but instead we meet more frequently depending on our changing schedules. While the exact time is not the same every week, we still meet 30 minutes - 2 hours every week depending on how much work needs to get done. In terms of group roles, Logan is still the leader and the primary communicator with our partner Dr Muñoz. But other than that, we have divided up the work each week between us equally depending on assignment tasks. Because we share responsibilities on the tasks and work, our roles are more loose and not restricted to the descriptions above. So far there have been no problems as we have been able to turn in assignments in a timely, through and complete manner.

14. Appendix A: Interview

Notes:

- Sea surface temperature between El Niño and La Niña
- Visualizations of Mississippi river data
- Time range vs average over time representations - average over time may be misleading because phases last for different periods of time
- No strong preference to transition between el nino and la nina, averaging across could be redundant
- 97-98 el nino 2015-2016 el nino, 1997-2000 monthly data - pick from 40 year window time slice, explore periods with floods
- Monthly slider linked to another vis on Mississippi delta
- Keep time the same in terms of Mississippi and El Niño links
- *Soil moisture*
  - creeps up/down during phases
  - bit more noisy than precipitation
  - Dr. Muñoz paper - mid 2009-2012 (potential prototype), 2018-2019 (most recent with clear flooding)
  - ENSO & soil moisture positively correlated
- warm pacific, high correlation/moisture; cold pacific, higher near lakes
- Soil moisture tuned to slider bar, as you move out of it it changes, more interactions w/ time periods
- 2011 flood might be best for seeing el nino to la nina and floods on missouri and mississippi river
- Can add different El Niño and La Niña events - example, show 2011 one then can compare to the 97 one
- Precipitation and river width/height
- Vicksburg, St. Louis, Memphis, daily data
- Time-series with river levels at different locations, pacific view, soil moisture view for our multiple views
- Start out with sea surface temp and slider bar, soil moisture data
- Any non geo spatial data vis :
  - Line graph of O/I index
- target users:

- researchers within the Muñoz lab, and potentially the readers of their work if the visualizations are insightful and the lab wants to distribute
- mission:
  - study the role of El Niño on regional climates, specifically the Mississippi river basin
  - convince policymakers and citizens to effectively respond to the challenges of flooding
- data analyzed:
  - weather, climate, and SST data from NOAA
  - soil, precipitation, and river level data from USGS
- user motives for analyzing the data:
  - better understand the critical role of El Niño on Mississippi flooding to allow for better disaster mitigation efforts.
- the possible insights the user is looking for in the data:
  - are there early warning signs for Mississippi river basin flooding, and if so, how can we prepare
- already completed visualizations:
  - the lab has created static visualizations of correlations between pacific SST and precipitation in the Mississippi river basin, but no interactive maps showing an explicit connection

15. Appendix B: Data Exploration

Data was gathered from NOAA and USGS. There are separate data sets for sea-surface temperature, soil moisture, precipitation, and river output.

Sea-surface Temperature (SST) [11]
- Time (year and month) **ordinal**
- Latitude **ordinal**
- Longitude **ordinal**
- Temperature (celsius) **quantitative**

Precipitation [12]
- Time (year and month) **ordinal**
- Latitude **ordinal**
- Longitude **ordinal**
- Height (centimeters) **quantitative**

Soil Moisture [13]
- Time (year and month) **ordinal**
- Latitude **ordinal**
- Longitude **ordinal**
- Surface Soil Wetness (ratio of calculated soil moisture (mm) to the maximum (760 mm)) **quantitative**

River Output [14]
- Time (year) **ordinal**
- Discharge (ft^3/s) **quantitative**

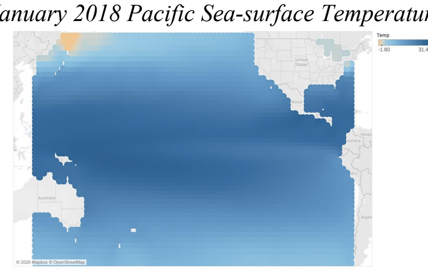
Data for each environmental condition obtained from various environmental sensors operated by NOAA and USGS and downloaded from their respective websites. Each source was downloaded in the form "netCDF", which is not a traditional format for direct use in Tableau or D3. Because of this, we converted the data to a csv format using a python script from the European Center for Medium-Range Weather Forecasts [15]. Once the data were in a flat-file form, we filtered the data to only include relevant times (2018-2019), latitudes (50S to 50N), and longitudes (70W to 130E).

Our data exploration process was quite informative, especially in elucidating the challenges that we must address as we fully flesh out our visualizations. For example, the Pacific SST map showed that it is difficult for a layperson to identify an El Niño event by simply looking at a map of water temperatures. Because of this, for our interactive SST temperature map with a slider across the El Niño spectrum, we decided that the final visualization should highlight, on the map itself, when a peak El Niño or La Niña occurs and what the bounds are for the SST anomaly. We also found that, at least in Tableau, the geospatial visualization of thousands of points can be slow, so we will have to experiment with limiting the number of points to make sure the interactivity is reasonably fast yet accurate.

In terms of trends, we were not able to determine yet whether the connection between ENSO anomalies and MRB environmental conditions is noticeable when visualized—this insight requires dynamic and linked maps, which we plan to utilize soon after learning brushing and linking techniques. We did, however, notice that the Pacific SST anomaly around the equator is easily visible in a given year, so we will be able to visualize this over time to easily view how the bounds of the ENSO anomaly change over time. When combined with an interactive slider that the user controls and highlights on the map at peaks, this will be a highly descriptive map that will be helpful in understanding ENSO.

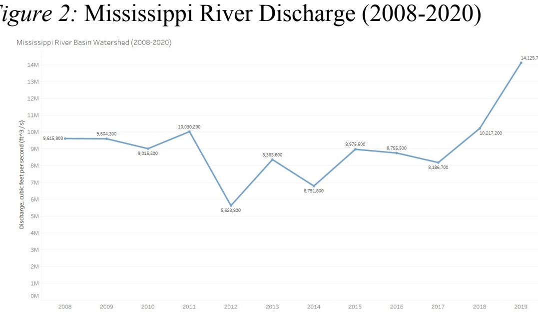
Notable Visualizations

Figure 1: January 2018 Pacific Sea-surface Temperature (C)



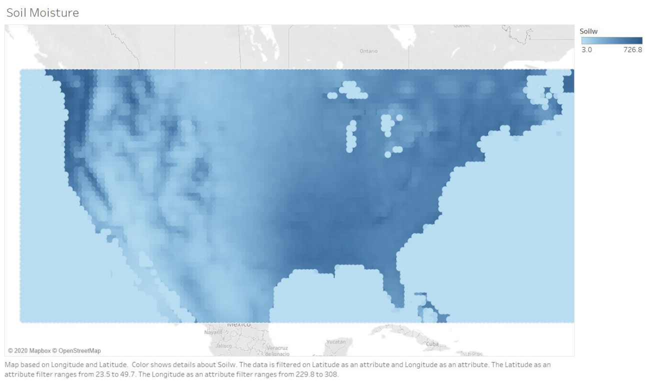
Since we were making static visualizations, we chose to visualize the starting year of the most recent full ENSO cycle, which will be the focus of our project. SST is a quantitative variable centered around 0, so we encoded the data using a color ramp with two hues (blue above 0, orange below 0) changing by spectrum to encode the data. It appears that in 2018, there is an SST anomaly extending from Ecuador about halfway across the Pacific. When we make the visualization dynamic and include multiple years/months, we will be able to see how that anomaly moves over time, which will be useful in understanding ENSO SST changes.

Figure 2: Mississippi River Discharge (2008-2020)



The data we explored with this visualization was the discharge of the Mississippi river near Vicksburg, Mississippi over 12 years. Because the data was a time series, we decided that a line graph would be best to show the differences between years and to see the overall increase in total Mississippi river discharge over the last few years. In order to clarify the exact discharge at any given year, we wrote the exact number right on the points in case the user would want an exact number.

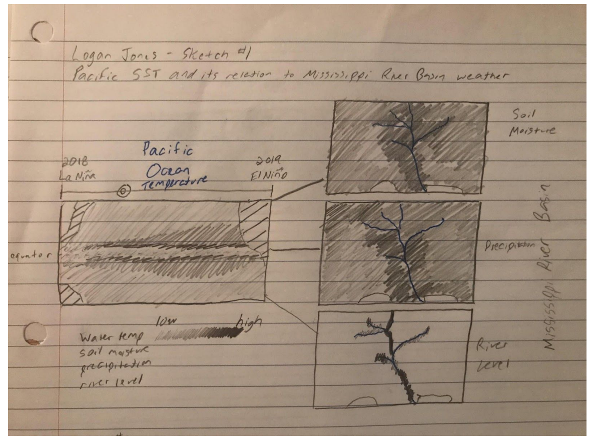
Figure 3: January 1 1948-2020 Average Soil Moisture in the United States (calculated surface wetness in mm)



One of the key variables to consider when studying ENSO and the MRB is soil moisture. A map was used to display geospatial data for the target areas (all within the US). There was almost 300MB of continuous data, so we decided to visualize monthly averages (still multiple megabytes of data) filtered for the new year in tandem with Figure 1's January SST. The averages for this period take into account data from January 1948 to September 2020 and the saturation of the mark (circles) represents the surface soil wetness dimension, measured in millimeters (maximum is 760mm) for each latitude longitude coordinate. From this data, we observe the Mississippi River Basin and Northeast coast have higher average soil moisture than the rest of the United States. Once we decide to use the entire continuous observed soil moisture data set in tandem with a time-series slider, we will be able to target distinct ENSO periods. We can then observe differences relative to the average soil moisture and compare this variable to others such as sea surface temperature and discharge. The sheer amount of data was surprising, and it may be challenging to convert the entire dataset into a readable format without partitioning it.

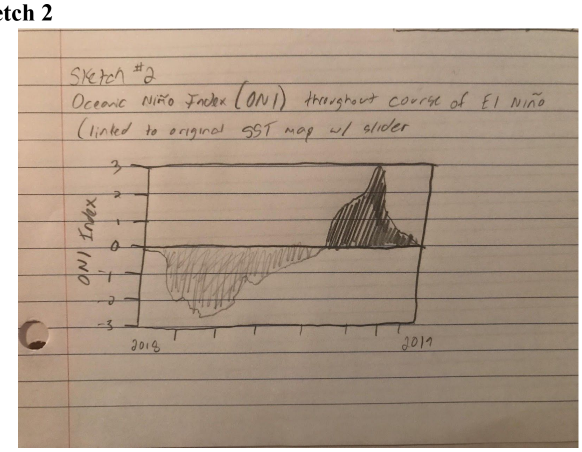
16. Appendix C: Design Sketches

Sketch 1



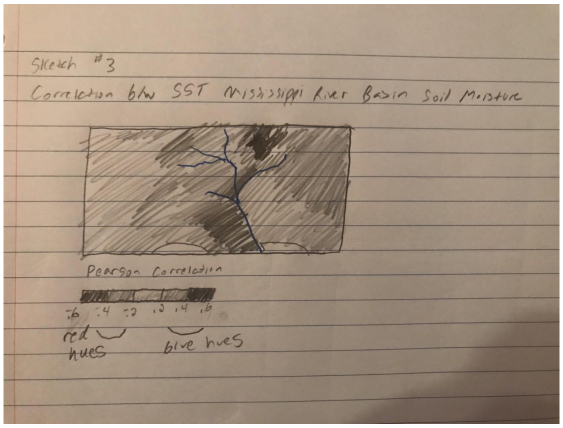
- **Favorite**
- Logan Jones
- In this visualization, I attempted to demonstrate the connection between ENSO anomalies and MRB conditions, which each of the three conditions in their own spaces. To encode environmental conditions in each map, I chose to use a color spectrum that changes by saturation since the data in each is quantitative. I chose points as the marks since the original data is based on latitude and longitude coordinates, which are best represented by points.
- This visualization addresses tasks 1-5 which are all related to viewing the environmental conditions of the Pacific and MRB and understanding their connection in the context of ENSO.

Sketch 2



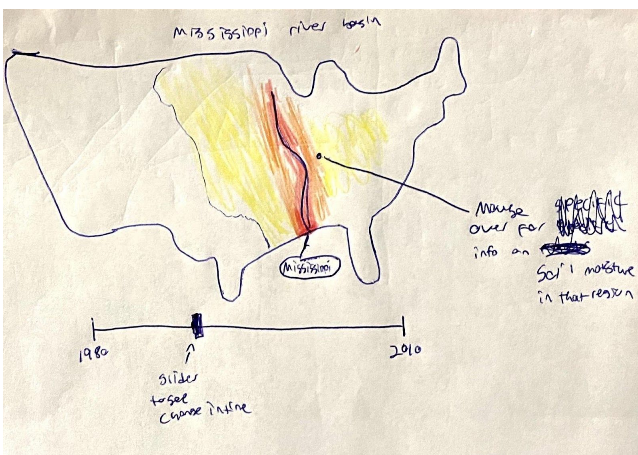
- Logan Jones
- In this visualization, I attempted to provide a non-geospatial interpretation of the ENSO spectrum. This line graph shows the Oceanic Niño Index (ONI) which represents the Pacific SST anomaly, over the period from 2018 - 2019 (the most recent El Niño cycle). The segment of the line below 0 represents colder temperatures associated with La Niña, and the segment above 0 represents the warmer temperatures associated with El Niño. Because ONI is ordinal around 0, I chose to encode with two different hues (above zero is red, below zero is blue) to show temperature differences. The line mark was most appropriate in this case given that I was graphing a single quantitative variable over time.
- This visualization addresses task 2, which is involved in viewing / understanding ENSO in terms of SSTs.

Sketch 3



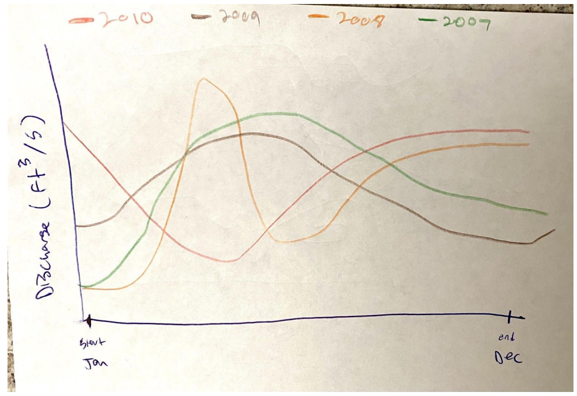
- Logan Jones
- In this visualization, I attempted to show the correlation between Pacific SST and MRB soil moisture, one of the variables known to be most closely related to Pacific El Niño conditions. The Pearson correlation coefficient is ordinal around 0, so it is conducive to a color spectrum with two hues, red (negative) and blue (positive). The data used for this visualization will be geospatial, so a map (area) is the most appropriate mark for this purpose.
- This visualization addresses task 1, which is exploring the connection between Pacific ENSO and MRB conditions.

Sketch 4



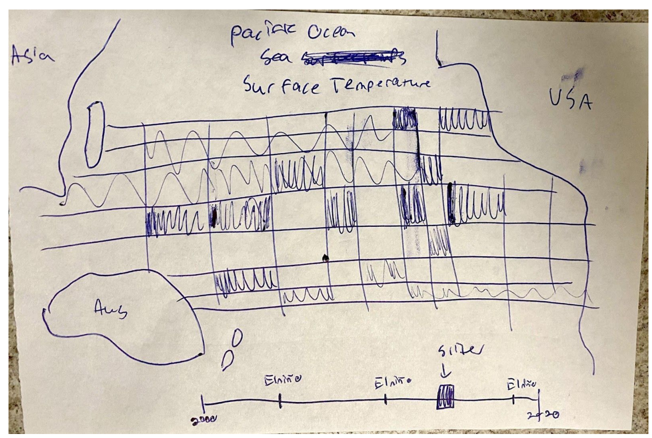
- **Favorite**
- Alexander Schad
- This visualization uses the area mark and the color channel to describe the soil moisture levels of the Mississippi river basin. Because it's a map showing the US, users would easily be able to see the highlighted Mississippi river and the surrounding area and compare soil moisture levels based on the color of each region. I included a slider that can be dragged by the user to change the time of the map. The color would be a converging scale where the darker red the area the more moist the soil would be. I believe color is a good channel for maps in seeing overall trends of soil moisture as the color should get darker the closer to the river a given point is. Also, hovering over a point on the map would enable the user to see specific soil moisture data based on that latitude and longitude.
- This visualization covers the 4th task: of comparing MRB soil moisture levels across the ENSO spectrum.

Sketch 5



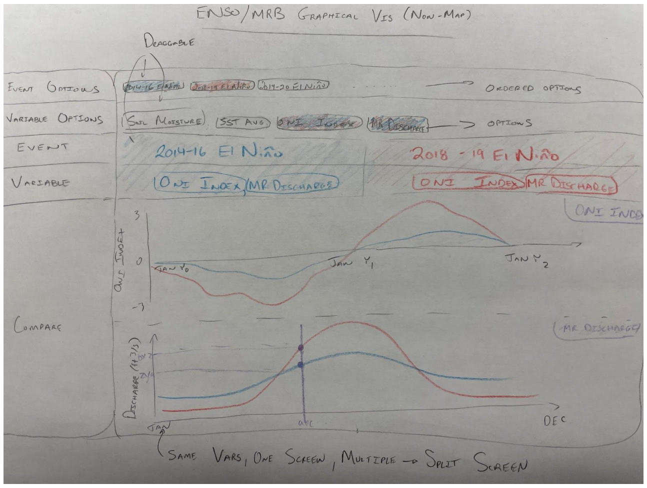
- Alexander Schad
- The marks I decided to use for this visualization were lines, and the channels are color and position. This visualization shows the years 2007 - 2010's Mississippi River Basin Discharge over each year where the x axis represents the months in the year and the y axis represents the discharge in ft^3/s. I used color to describe the years because time is ordered data but can be viewed as categorical. Position is best to show relative differences in discharge based on the given time.
- This sketch covers the 5th task of comparing MRB river levels across the ENSO spectrum.

Sketch 6



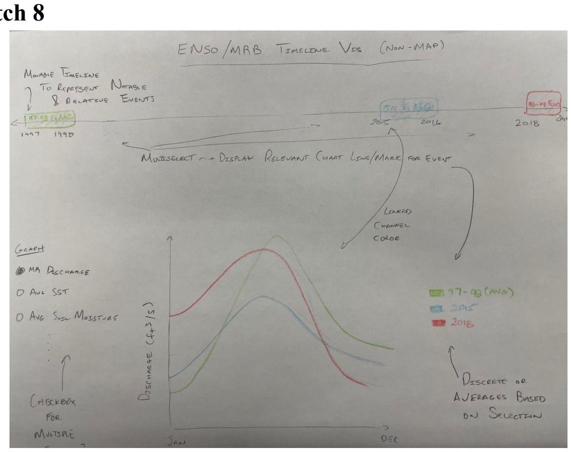
- Alexander Schad
- This simple sketch shows a heatmap of the Pacific Ocean sea surface temperature. The darker the shaded in square is, the higher the temperature is in that region. I included a slider at the bottom to enable the user to see the map at various years. The El Nino events would be marked on the slider so users can compare SST during El nino and when it's not El nino. I chose color as a channel and area as a mark because I believe that seeing the differences in shading between regions will enable users to clearly see differences in sea surface temperatures.
- This visualization would address task 2 in viewing sea surface temperatures during ENSO cycles.

Sketch 7



- Sina Soltanieh
- The marks used for the main chart views are lines, and the channels are color and position (vertical + horizontal). This visualization represents a selection of variables that can be indicators of ENSO conditions, and a list of ENSO and high flooding periods, which can then be combined into comparative views (as it is difficult to distinguish between normal periods and ENSO in many time series visualizations, especially within just juxtaposed maps). In this particular case, the ONI Index and Mississippi River discharge are being compared for the 2015-16 El Niño and the 2018-19 high-flooding El Niño.
- This vis sketch covers tasks 3 through 5, where we are interested in comparing different variable indicators across the ENSO spectrum.

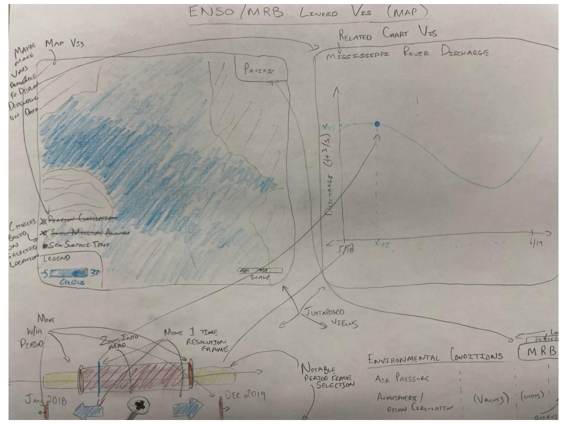
Sketch 8



- Sina Soltanieh
- The marks used for the main chart views are lines, and the channels are color and position (vertical + horizontal). This visualization represents a timeline of notable ENSO and high flooding periods, which can then be combined into multi-select comparative views with color channels as the dimension to categorically distinguish these periods (as it is difficult to distinguish between normal periods and ENSO in many time series visualizations, especially within just juxtaposed maps, so continuous may be difficult). The relevant value can then be selected to generate the proper charts (heatmap of SST, moil moisture, MR flooding, etc.)..

- This vis sketch also covers tasks 3 through 5, where we are interested in comparing different variable indicators across the ENSO spectrum.

Sketch 9



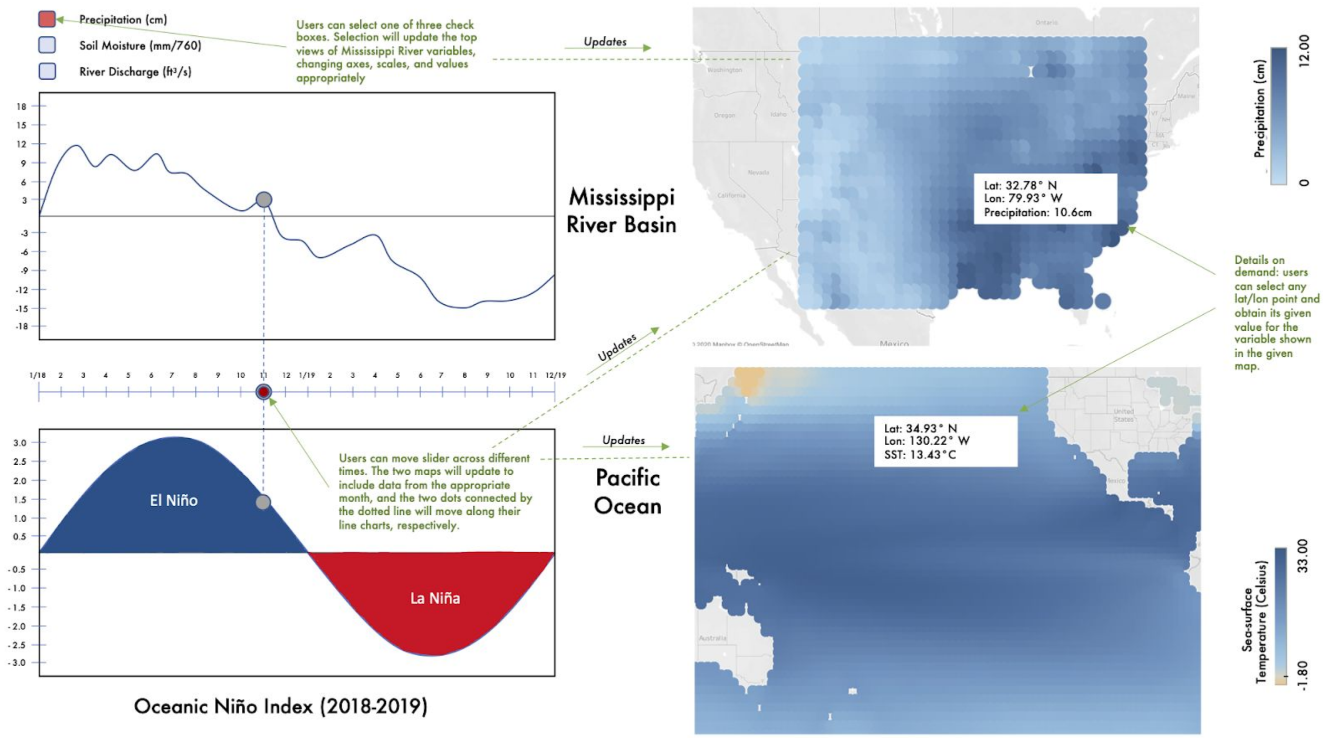
- **Favorite**
- Sina Soltanieh
- The main point of differentiation this visualization provides is a slider bar, which can be tuned to specific ENSO periods, and further slid or moved by frame of time resolution via arrow buttons or keys (represents the time dimension). The bottom right view presents a conglomerate of environmental condition info we would want to present for the zoomed in area, if any, as well as tooltips to provide descriptions. The main views are juxtaposed on the top of the screen. In this example, the map view is showing sea surface temperature in the pacific, with hue as the quantitative channel representing temperature in celsius, and latitude and longitude dimensions as the map points. The graph view to the right shows Mississippi River discharge over the selected time interval from the slider. These variables can then be adjusted to try to connect geospatial data to time-series data such as the ONI Index, MR Discharge, or any other indicator we would like to support..
- This visualization sketch covers tasks 1 through 6, all of our desired tasks. We can compare different variable indicators across the ENSO spectrum, view pacific SST and environmental tasks, and most importantly, it covers our highest priority task of exploring the connection between ENSO and MRB conditions.

Sketch #1, which shows all four environmental variables connected in four separate maps, is a central visualization to our project as it highlights the connection between ENSO conditions (SST) and MRB conditions (soil moisture, precipitation, and river output/level). The slider bar allows users to explore this connection over the entirety of the most recent ENSO cycle (2018-2019). Sketch #4 shows the soil moisture levels of the surrounding Mississippi River Basin. The slider allows the user to see the change in levels over time. While the visualization gives a summary of the soil moisture data in the MRB, users can interact with the visualization by moving the slider and hovering over a specific point on the map. This interactivity promotes discovery from the users which will help their scientific findings and conclusions. Sketch #9 shows Pacific sea-surface temperatures in the map view and Mississippi River discharge in the graphical view, which are linked together by the selected time frame or

ENSO period. There are many interactions to help the user in data comparison and exploration, such as zooming into a location updating environmental variables and time adjustments updating both the map and graph. However, we do not want too many interactions and views, as this may impose a relatively high cognitive load, and subsequently make it more difficult for users to explore trends and compare the correct variables they are looking for.

Our previous assignments made it clear that a major challenge in our project is simultaneously showing these variables and the connections between them over time. Mapping four separate variables on one map was out of the question given the drawbacks of superposition, so we sought to find a layout of the four maps that was most conducive to showing the ENSO / MRB connection in an aesthetically pleasing manner. For Sketch #9, we tried to address this challenge by partitioning variable representations into both graphical and map views. Variables such as Mississippi River Discharge, which is best represented using the line chart visual encoding, or relevant indices to display a correlation between ENSO related variables can be represented without a map, while sea-surface temperature and other map based data that depends on neighboring values to create a colormap (the most effective encoding for the data) would best be represented via a map view. Color considerations are also of great importance. For Sketch #4, the color channel uses a converging luminance scale that is best for quantitative data such as soil moisture. When dealing with the area mark, color pops out as a channel because position is used up inherently with the shape of the map, so color is the next best option. Hue doesn’t work for the color scale because hue doesn’t have an inherent ordering whereas differences in luminance can be easily identified by the human brain. The final challenge we are constantly looking to address and maintain is how to best present the data in a scientific and unbiased manner, which is very important to the ethics aspect of data visualization. By focusing more on trend exploration with scientific data rather than the presentation of specific research findings, we have already made a big step, and we are always thinking about how to best remain professional, scientific, and unbiased in our data visualization project.

17. Appendix D: Digital Sketches



The visualization above addresses all 6 of the tasks we outlined in the task analysis table above. The visualization also features a graph showing the Oceanic Niño Index over time. This feature was not included in our original task analysis table so we added it as a 7th task. In terms of which tasks are prioritized for this digital sketch, each of the four visualizations are given an equal amount of space on the screen. Because we are showing two graphs and two spatial maps at the same time, none are given more precedence over the other. Our primary task is exploring the connection between ENSO and MRB conditions which is accomplished by comparing the ENSO cycle represented by the bottom two visualizations and the MRB conditions represented by the top two. Tasks 3-5 can be accomplished by choosing whichever variable you want at the top which will change the top left graph accordingly. Task 6 is accomplished by using the time slider at the middle left of the visualization to compare the graphs change over time. Task 2 can be accomplished by looking at the bottom right map and task 7 is accomplished by looking at the bottom left graph. By having these four visualizations present to the user all at once, they can easily compare data from one quadrant to another. Since our high level task is discovery, having 4 visualizations that are all connected enable users to test and confirm hypotheses. Everything else has stayed the same from our previous plan as we feel that this visualization meets all of the project guidelines and meets our partner's expectations.

18. Appendix E: Reflections

Logan

I have been very impressed with our communication throughout this project. Our original plan was to meet once per week to plan and execute work, but we came to realize that meeting more than once was often necessary and planning a mutually applicable meeting time was often difficult. Luckily, all three of us were very flexible and willing to meet as much as needed to get work done well. We have been communicating via text to set up meetings or quickly discuss our project and via Zoom for our lengthier discussion and collaboration sessions. Our communication has been most effective in the sense that everyone has been timely, professional, and highly collaborative. One area that could be improved is in our collaboration via Github. At least in my case, I am new to using Github and it has been a challenge for me to learn how to resolve merge conflicts, but luckily, I have gotten more comfortable with this over time.

Sina

Communication with the group went beyond my expectations. Setting up appointments with Dr. Muñoz via email sped up primary communications, and group messaging worked wonders in quickly scheduling meetings with our group. Everyone was responsive and we were able to set up multiple meetings a week compared to our original scheduled one meeting per week to account for the increased amount of work required for each sprint. We also did a great job of outlining assigning tasks for each sprint for the project, paper, and video. Zoom scheduled meetings made it easy to reconnect, and Logan sent out email notes after meetings highlighting what has been completed, and what tasks remained for each person in the group. It was nice having done a software engineering co-op so I could help my group mates with technical problems such as git merge conflicts in a relatively efficient manner. Overall a very professional experience with regards to communications.

Alexander

Throughout our whole project, we have been able to communicate effectively and timely in order to meet deadlines and to accommodate our own personal schedules. We have a text group chat where we communicate when it is necessary to go over the plan for the week or any tasks that need to be addressed. Originally, we planned to meet once a week every Thursday morning, but we ended up needing to work more than once a week for some weeks and sometimes Thursday morning was not the most efficient time to meet so we would meet on another day or time. During our zoom meetings, we would go over work that needed to be completed and when we should complete it. This was also written down by Logan and emailed to us after the meeting so we could look back and see tasks that needed to be done during the week. Our communication with Dr Muñoz was very efficient as well considering whenever we had a question about the project data or about the visualizations, he would provide an answer very quickly. In terms of what could be improved upon, sometimes I would have technical difficulties a few times when we tried to meet. This would result in me showing up to the meeting late and to solve this I ended up joining on my phone when my computer wasn't working. Also, being in my second year having less experience than both Logan and Sina, it felt nice being able to ask either of them for help if I didn't understand how to code something. All in all, our communication was very instructive and thorough to enable us to work together more efficiently throughout the project.

19. Appendix F: Slides

Project Slides