8-2022

Network Analysis of Scientific Research in the Gulf of Mexico

Juliet S. Vallejo
The University of Texas Rio Grande Valley

NETWORK ANALYSIS OF SCIENTIFIC RESEARCH IN THE GULF OF MEXICO

A Thesis

by

JULIET S. VALLEJO

Submitted in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Major Subject: Ocean, Coastal, and Earth Sciences

The University of Texas Rio Grande Valley

August 2022

NETWORK ANALYSIS OF SCIENTIFIC RESEARCH IN THE GULF OF MEXICO

A Thesis
by
JULIET S. VALLEJO

COMMITTEE MEMBERS

Dr. Owen Temby
Chair of Committee

Dr. Dongkyu Kim
Committee Member

Dr. Andrew M. Song
Committee Member

Dr. Christopher Gabler
Committee Member

August 2022

ABSTRACT

Vallejo, Juliet S., Network Analysis of Scientific Research in the Gulf of Mexico. Master of Science (MS), August 2022, 58 pp., 9 tables, 10 figures, references, 66 titles.

The Gulf of Mexico ecosystem represents a significant management challenge in its 3,423-mile coastline and transnational regulatory setting, along five states from the United States, six states from Mexico, and Cuba which borders the southeastern quadrant. Providing various resources to the regional economies, the continued success of these resources depends on the collaboration among transnational participants in bringing together complementary skills and multidisciplinary approaches to producing, circulating, and utilizing scientific knowledge. Using bibliometric analysis of Gulf of Mexico-related published research over 18 years from 2000 to 2018 allows for identifying organizations, their connections, and trends in the production of scientific research about the region. The results reveal a robust network structure between government and academic institutions but a disjuncture in US-Mexico cross-border research, with organizations outside of the Gulf of Mexico region having a stronger relationship with institutions in each country. The database and findings provide potential information that can contribute to the ongoing efforts to improve transboundary collaboration in the Gulf of Mexico region.

DEDICATION

The competition for my master's studies would not have been possible without the love and support of my wonderful family and friends. Through the long hours and unforeseeable challenges, they have been by my side and have provided me with guidance and encouragement. Thank you for all your help, love, and patience.

ACKNOWLEDGMENTS

I am incredibly grateful to Dr. Owen Temby, chair of my thesis committee, for all his advice and assistance. From funding, research design, and manuscript editing, he encouraged me to complete this process through his infinite patience and guidance. My thanks also go to my thesis committee members: Dr. Dongkyu Kim, Dr. Andrew Song, and Dr. Christopher Gabler, who provided feedback and guided me through research design and data processing, ensuring this work's quality.

Additionally, I would like to acknowledge the staff at the Center of Survey Research and Policy Analysis at the University of Texas Rio Grande Valley, who assisted in the data management portion of my research, ensuring that the dataset became more manageable in finalizing the analysis for this paper.

I also want to thank my colleagues: Katia Sanchez, Derek Katznelson, Evelyn Roozee, and Mya Brown, for their support. Finally, I would like to acknowledge the NOAA Center for Coastal and Marine Ecosystems and the Education Partnership Program for supporting my studies and this research.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The Gulf of Mexico region is a rich and diverse ecosystem that provides vital services such as oil and gas production, tourism, habitats for various species, and support for local economies. It is a deep, semi-enclosed oceanic coastal basin whose shores border the United States, Mexico, and Cuba, with a tight connection to the Atlantic Ocean through the Straits of Florida and the Yucatan Channel. The varying biodiversity in its marine habitats includes estuaries, wetlands, shallow inshore waters, reef communities, and deep-sea, providing an estimated 14% of commercial and 30% of recreation U.S. fishing revenue in 2020 (United States. National Marine Fisheries Service, 2020; United States. National Ocean Service, 2011; Zaldivar-Jimenez, Ladron-de-Guevara-Porras, Perez-Caballos, Diaz-Mondragon, & Rosado-Solorzano, 2017). Adding the three other significant industries in the Gulf of Mexico, tourism, shipping, and petroleum/natural gas production account for an estimated total of $100 billion in 2013, 17% of the total U.S. Gross Domestic Product (GDP) (Alvarez Torrez, Rabalais, Pina Gutierrez, & Pardon Lopez, 2017; United States. National Ocean Service, 2017; United States. National Ocean Service, 2011). While one of the significant economic activities in Mexico also consists of petroleum/natural gas production, with an estimated $50 million in 2018, 2.11% of the total Mexico GDP (Banco de Mexico, 2018). Resources of the gulf are essential to the economy of the communities in the region, yet, the Gulf of Mexico is under constant stress by environmental alterations to the coastal ecosystems, including climate change, extensive shoreline development, water engineering systems, exposure to significant storms, or coastal wetland degradation (Alvarez Torrez, Rabalais, Pina Gutierrez, & Pardon Lopez, 2017). Vulnerable especially to changing weather systems or environmental hazards, where hurricanes and drought have significantly impacted not only economic security but human health, safety, and overall habitat of the gulf.

The challenge in managing a complex system such as the Gulf of Mexico is not only in the wide variety of habitats but the transboundary regulatory setting between Mexico, the United States, and Cuba adds a social, political, and administrative aspect to the system. As such, resource managers and policymakers must balance scientific information, address their community's different needs and interests, and maintain a robust, resilient, and productive ecosystem on which services can be depended in the long term (Djenontin & Meadow, 2018). The importance of acknowledging these links between science and society within this ecosystem has increased the investment of government and regional stakeholders in scientific research to clarify any issues and create appropriate responses in decision-making (Cruz & McLaughlin, 2008; Delacamara, O'Higgins, Lago, & Langhans, 2020; Sherman, 2014). This increase has led to the interest in a co-production approach in the production of scientific knowledge.

CHAPTER II

REVIEW OF THE LITERATURE

The approach to creating co-production knowledge appeals to many, compared to the traditional linear network approach, mainly due to the context-based, pluralistic, goal-oriented, and interactive principles. Not only increasing the relevance and usability of any findings, but this innovation allows progress to occur no matter how complex (Djenontin & Meadow, 2018; Nostrom, et al., 2020). The critical need for scientific knowledge and interest in the co-production approach has led to stakeholders' creation, support, and direct involvement in programs. An example of one such program is the NOAA RESTORE Science Project, a governmental program through a U.S. scientific and regulatory agency, the National Oceanic and Atmospheric Administration (NOAA). This program provides funding and a partnership to scholars from various institutions with a U.S. scientific and regulatory agency to construct and monitor scientific projects based on the long-term sustainability of the resources in the Gulf of Mexico (United States. National Centers for Coastal Ocean Science, 2022). This project allows for high-quality knowledge to be co-produced and available for quicker distribution to those interested. Another method by which stakeholders have approached the efforts in the co-production of knowledge is through the foundation of collaborative partnerships, a network of stakeholders dedicated to working on common issues. An example of a collaboration partnership effectively working to solve complex problems is the Sacramento River Science Partnership. A voluntary joint enterprise consisting of different stakeholder groups of different fields whose mission is to develop, disclose, and discuss management efforts on the resources found in California's Sacramento River (Vallejo, 2021; Sacramento River Science Partnership, 2021). The success of such partnerships has shown that the use of network arrangements in co-production allows for research to be created and spread over interested audiences rather than if individual organizations worked on the issue alone (Gunton, Peter , & Day, 2007; Conley & Moote, 2003; Lubell, 2014).

An effective communication network between stakeholders must be established for any form of co-production and effective management to last long term (Beier, Hansen, Helbrecht, & Behar, 2017; Cinner & Barnes, 2019). Other long-term collaborative partnerships have been shown effective through transboundary co-production, such as one between the United States and Canada in the Great Lakes region (Song et al., 2017). In contrast, one of the most prominent collaborative partnerships identified in the Gulf of Mexico region, the Gulf of Mexico Alliance (GoMA), seems to lack that transboundary nature. Formed in 2004, the goal of GoMA is the improvement of the ecological and economic health of the Gulf of Mexico region through providing resource management tools and mechanisms of coordination to collaboration (GoMA, Gulf of Mexico Alliance, 2019; Wondolleck & Yaffee, Mobilizing a Multistate Partnership in the Gulf of Mexico, 2017). While having an active presence in the United States in a cross-state capacity, its involvement as a transboundary organization has not been evidenced.

Establishing transboundary collaborative partnerships would facilitate the co-production of complex research, which not only assists in framing problems but in informing responses, stimulating public debates, and monitoring the effectiveness of policy actions (Wondolleck & Yaffee, 2017). Critical for effectively implementing measures in sustaining long-term fiscal and environmental accountability of the Gulf of Mexico resources and the communities that depend on them. However, essential questions must be answered to understand the underlying context of innovation in the governance systems among these countries. First, we must consider which organizations comprise the scientific knowledge production that supports the Gulf of Mexico and to what extent co-production occurs across organizational, institutional, geographical, and jurisdictional boundaries. In exploring the production of research based on the Gulf of Mexico through bibliometric network analysis, a statistical method analyzing published research, we will get a preliminary look into the elements of co-production and distribution of scientific knowledge and the previously unclear trends. Bibliometric network analysis

The analysis of published research is a prevalent method to uncover emerging trends in articles and journals, identify research institutions, or uncover network relationships. While the application of this method is relatively new, the advancement, availability, accessibility, and multi-disciplinary nature of this analysis have allowed for its use in various fields (Wasserman & Faust, 1994; Van Den Besselaar & Sandstrom, 2019). The flexibility of applying this method is even evident in the basic procedure of bibliometric analysis: (1) defining the aims and scope of the study, (2) choosing the techniques to use for the analysis, (3) collecting the data, and (4) run the analysis and report the findings (Donthu, Kumar, Mukherjee, Pandey, & Lim, 2021). However, the systematic and replicable way of bibliometric analysis allows for the comparison of any information collected (Milesi, Brown, & Schneider, 2016). Several papers have even shown the use of authorships and the co-production network in the management field, either by the use of co-citation analysis which focuses more on the intellectual structure of management, or as the one used in this paper conducted using co-authorship analysis (Koseoglu, 2016; Ramos-

Rodriguez & Ruiz-Navarro, 2004; Miranda-Gonzalez, Aref, Theile, & Zagheni, 2020). The application of co-authorship analysis and the network structure found from using it inferred an evolution through the formation and incorporation of knowledge, the consolidation, and dissemination to the various stakeholders, and the transformation to better expand into areas where trends show a lack (Glänzel & Schubert, 2004). The application of quantitative techniques in bibliometric analysis can be divided into two categories: performance analysis, which accounts for the contributions of research with a myriad of ways to measure, and science mapping, which focuses on the network between contributors and obtaining the visualization of the network analysis (Donthu, Kumar, Mukherjee, Pandey, & Lim, 2021; Narin & Hamilton, 1996; Noyons, Moed, & Van Raan, 1999; Van Eck & Waltman, 2010).

The data often used in this form of analysis tends to be significant in volume, running in the hundreds if not thousands, and allows for objective and subjective evaluations to decipher and map the unstructured data into something that makes sense (Rogers, Szomszor, & Adams, 2020). Therefore, allowing for the results of these bibliometric studies to form foundations for advancing challenges in new, meaningful ways where scholars can have a better overview of a field, identify gaps in knowledge, inspire new ideas for research, and better manage contributions. Many programs are available to assist in assisting in analyzing bibliometric data, from VOSviewer, Gephi, Bibliometrix, CitSpace, and Pajek, to packages that can be added to analytical software such as R Studio, BibExcel, or more commands to input (Bornmann & Ozimek, 2012; Donthu, Kumar, Mukherjee, Pandey, & Lim, 2021) . In the case of the methods VOSviewer, the computer programming software many scholars use, its algorithm uses a combination of two formulations. First is modularity-based clustering, a variant of the clustering algorithm used to detect communities in a network and measure the quality of the network structures (Newman, 2001; Van Eck & Waltman, Visualizing bibliometrics networks, 2014; Su, Peng, & Li, 2021). At the same time, the multidimensional scaling method is used to visualize the level of similarities and distances between the different objects or nodes being analyzed (Van Eck & Waltman, Software survey: VOSviewer, a computer program for bibliometric mapping, 2010). The mapping technique used by the VOSviewer program uses a co-occurrence matrix that measures for normalizing co-occurrence data through a measure known as association strength or proximity index; a network is then created by the finite set of nodes and links between them.

Network analysis is a theoretical perspective and set of techniques used to understand and qualitatively measure relationships, the set of notes, and their links. It is possible to identify the most critical nodes then. The choice of nodes to be analyzed depends on the context in which a given data set is inserted. These nodes can be individuals, groups, organizations, and even countries; each relationship defines a different type of network (Bennett & Gadlin, 2012; Haythornthwaite, 1996; Valente, 2010). From the perspective of network analysis, the relations are not properties of nodes but whole systems.

Case Study: Great Lakes of North America

The use of co-authorship of published papers has been widely used as a proxy to measure scientific co-production between different institutions (Adams, Black, Clemmons, & Stephan, 2005; Belter, 2013; Cavadas, 2020; Elango & Rajendran, 2012; Gunton, Peter , & Day, 2007). One such study is the bibliometric paper "Assessing transboundary scientific collaboration in the Great Lakes of North America” by Andrew Song. Offering an initial assessment of the Great

Lake scientific co-production by identifying which organizations are part of the production of scientific research and to what extent they collaborate across organizational, institutional, and jurisdictional lines in efforts to better the transboundary governance between the United States and Canada. The findings showed a robust relationship between government and educational institutions. However, a trend in the disjuncture between USA-Canada cross-border collaboration was observed, even with the transboundary nature of social-ecological challenges. These findings promote a more significant effort in bridging the gap between the two countries and ultimately improve the integration and efficiency of knowledge production through collaboration.

Research Aims

An initial assessment of scientific co-production networks using published articles in combination with statistical analysis will assist in identifying leading stakeholders in scientific research and create a visual representation of network trends within the Gulf of Mexico region, which are currently unknown. Using the methodology from the Great Lakes paper, the results presented in this research is divided into three key findings: (1) publication outputs, journals, and inter-organizational collaboration, (2) geographical and inter-jurisdictional collaboration, and (3) inter-institutional collaboration.

CHAPTER III

METHODOLOGY AND FINDINGS

Methodology

In the research, bibliometric network analysis uses statistical methods with published papers by co-authors as a proxy to measure collaboration between institutions. The methods used in this study follow established contours of assumption, procedures, and caveats previously used in similar types of analyses of a transboundary region (Song, Hickey, Temby, & Krantzberg, 2016). The data is downloaded from a reputable bibliographic database, followed by identifying, coding, and manipulating the variables needed for analysis using data management software. In visually representing the networks from the data in a bibliometric map, dyads were created from the variables showing the strength and weight of the links created.

Initial considerations

One of the databases recognized to contain the most comprehensible and reliable sources of published research across various fields is Elsevier's SCOPUS database. Two other recognized database is Thompson Reuters' Web of Science (WoS), the first broad-scope bibliographic database created, and the newest, most convenient database of Google Scholar. While each has its benefits as a tool for scholars, it has been found that not all indexed journals can be found in comparison to the SCOPUS database (Mongeon & Paul-Hus, 2015; Powell & Peterson, 2017; Kalhor, Sarijalou, Sadr, & Bahrak, 2022). Even so, there is a limitation in using either of these databases, which consists of the biases favoring Natural Sciences and Engineering research publications over Social and Art Sciences and the overrepresentation of English-language journals over other languages (Harzing & Alkangas, 2016; Mongeon & Paul-Hus, 2015; Pranckute, 2021; Vera-Baceta, Thelwall, & Kousha, 2019). There must be caution in interpreting the data as a comparative evaluation in transnational research systems. However, the benefits of a practical, cost-effective, verifiable, and discreet method outweigh the drawbacks of the overall research.

In using the similar methodology process from the Great Lakes Research (Song, Hickey, Temby, & Krantzberg, 2016), where they manually processed their data and used an adjacency matrix to compute all pairs as the method of analysis, that is going to be changed in the research for the Gulf of Mexico. This change is primarily due to the overall amount of data processed being double what was produced in the Great Lakes. Two data programs were chosen to decide which software program would be best suitable for this project. The widely used Microsoft Excel was primarily used for primary data manipulations. Two software programs were used in the data management process: Microsoft Excel and the general-purpose statistical software STATA by StataCorp. Using these two programs is the ease with which each does different jobs. While Microsoft Excel is excellent for its accessibility and user-friendly organization, editing, basic calculations, and visualization methods, STATA is powerful as an analytical program that can quickly process large datasets and manipulate data. We can obtain a clearer understanding of the bibliometric data using these two programs in combination with the network visualization software program VOSviewer.

Search procedures

Data were obtained from the SCOPUS database on August 14, 2019, where the search criteria were compromised into two parts. The "thematic" which focused on Gulf of Mexico science as it relates to the rounded and interconnected view of the system (e.g., search terms such as "ecosystem," "watershed/way," "climate," "policy," and "management") This was combined with a "geographic" criteria (e.g., "Gulf of Mexico," "North America," "Mexico," and "Cuba") to search for papers that met both criteria in either the title or abstract of the journal. The search focused on 18 years from 2000 to 2018 and compromised articles, including notes, reviews, and editorials.

```
(TITLE-ABS ( ecosystem OR water* OR ecolog* OR biodiversity OR climate OR policy OR
governance OR management ) )

AND

( TITLE-ABS ( "Gulf of Mexico" AND ( "North America" OR "North American" OR mexico
OR mexican OR us OR usa OR "United States" OR texas OR louisiana OR alabama OR
florida OR mississippi OR tamaulipas OR veracruz OR tabasco OR campeche OR
yucatán OR "Quintana Roo" OR "Atlantic Ocean" ) ) ) AND ( ( PUBYEAR = 2018 OR
PUBYEAR = 2017 OR PUBYEAR = 2016 OR PUBYEAR = 2015 OR PUBYEAR = 2014 OR
PUBYEAR = 2013 OR PUBYEAR = 2012 OR PUBYEAR = 2011 OR PUBYEAR = 2010 OR
PUBYEAR = 2009 OR PUBYEAR = 2008 OR PUBYEAR = 2007 OR PUBYEAR = 2006 OR
PUBYEAR = 2005 OR PUBYEAR = 2004 OR PUBYEAR = 2003 OR PUBYEAR = 2002 OR
PUBYEAR = 2001 OR PUBYEAR = 2000))
```

Figure 1: Search terms used in the SCOPUS database.

Due to the number of journals found matching the search criteria in SCOPUS, the files downloaded had to be obtained in 12 separate spreadsheet files and later merged into one file using Excel software. The data obtained included author name(s), addresses, article title, year of publication, abstract, and journal title. A total of 8,001 papers met the selection criteria, but an initial manual quality control had to be applied to check for relevance and duplication. The majority of the 3,398 articles excluded in the preliminary round of quality control consisted of single-author articles. An observation made as these articles identified is that they were primarily published in the energy/engineering field of study. A final set of 4,603 published articles from the original set were kept for further analysis in this study.

Data management

The data obtained from SCOPUS is categorized into columns: year, journal, article title, author(s) name, and author(s) address. To correctly identify the different networks occurring within the data, we had to separate each address for each author. We used this information to code the different organizations and locations to correctly classify each organization according to its political jurisdiction and institution type.

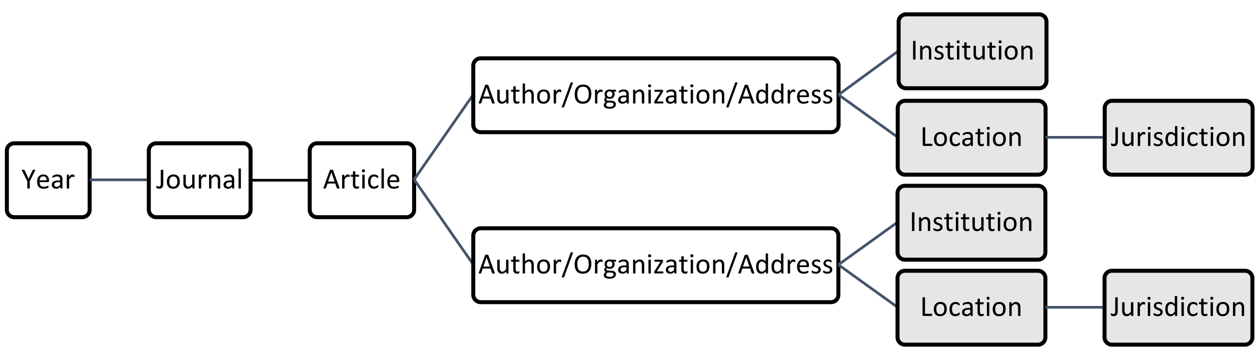

Figure 2: Data identification per article

Records were downloaded into an Excel program combined with STATA, and each separate author was transposed into rows containing all their data in separate columns. In collaboration with the Center of Survey Research and Policy Analysis staff at the University of

Texas Rio Grande Valley, the manual coding was completed to identify the different organizations, institution types, geographical locations, and jurisdictions. In this process, organizations were rewritten uniformly, as the original data often contained different spelling variations of the same organization. Using the organizations found in the dataset, the classification of institution type was identified. The author's address was used to identify the country each organization is located in, and if it found either the United States or Mexico, further classification to identify the state was also done. This resulted in the identifying 19 jurisdictional categories: U.S. National, MX National, Cuba National, country of Cuba, 11 Gulf of Mexico states, U.S. states outside of the Gulf of Mexico region, MX states outside of the Gulf of Mexico region, international locations, and inter-governmental entities [e.g., United Nations Environment Program]. For institution types, we coded 20 categories: Inter-governmental agencies, US-based set of national, sub-national [e.g., state or municipality agencies], education, non-government/not-for-profit (NPO), for-profit (private), and the equivalent set for Mexico, Cuba, and international institutions.

Using the STATA program coding [view Appendix – STATA.do file: Data Management] further quality control followed in finding if two or more authors in the paper were from identical organizations, pairings were not considered inter-organizational collaboration and were omitted from further analysis. Starting with 15,383 authors, removing duplication or single authors brought it down to 13,067. We did encounter one outlier; an article by *the Bulletin of the American Meteorological Society* was identified to have 228 authors listed from all over the world. After reviewing the article and much consideration, this article was removed. The reasoning behind this decision is that after running the analysis both with and without the outlier, it was found that it created relationships where there is none if excluded affecting the overall network centrality and bringing the final number of authors to 13,065 (Mihiri Shashikala, George, & Shujaee, 2015). Following this, we created I.D. for each variable to analyze. Identifying 2,490 organizations, 82 countries, and 160 locations accounts for the different states in the United States, Mexico, and international countries.

Analysis and visualization procedures

In analyzing publication outputs, journals, and inter-organizational collaboration, we had to create a new file formatting the raw data into a format allowing comparisons to the collaboration data. Using STATA program coding [view Appendix-STATA.do file: Publication Analysis], we obtained the number of journal publications and articles per year for the original and collaboration data. In the raw data, we identified 1,648 publishing journals and 7,948 articles compared to the 1,059 publishing journals and 4,245 articles identified to participate in collaborative research. Additionally, the frequency of organizations found in each variable was calculated. The information obtained by this portion of the analysis was compiled into a single file where tables and graphs were created to display the results. When looking at the data obtained from the.do files, many of the steps and files created are not in the final presentation of the data due to only using them to check the validity of the calculations through randomized observations.

To obtain the total number of observed pairs within each article, we created all pairs within groups, or dyads, in STATA [view Appendix – Stata.do file: Network Analysis]. The approach used the 'joinby' command, where the data was saved in a temporary file, and the variable names were changed before using the command. The code fragment created all possible pairs without duplicate permutations, amounting to7,515 pairs (UCLA: Statistical Consulting Group, 2021). Additionally, to find the states within the Gulf of Mexico geographical focus, we repeated the process by excluding states and countries outside the region. This process produced a total of 9,024 pairs.

In analyzing each variable's network data, two separate files were created for the various variables in a format allowing the VOSviewer program to read and output results with a map representing the network (Van Eck & Waltman, VOSviewer - Getting Started, 2022). The first file is a node list or map file containing information about the graph's variables. These variables are characterized by several attributes, each corresponding to its column in the file. The second file is an edge list, or network file, containing information about the links between the variables in the graph, which defines which pair of variables are connected by a link and the strength of each link. The files input into VOSviewer is in the sparse format, where three columns were created and saved into a comma-separated value (.CSV) file format. The first two columns specify the assigned I.D. of pairs connected by a link, the first column containing the source I.D. while the second includes the target I.D. The third column specifies the strength of the link, which is the number of repetitions of the exact pair of I.D.s. The edge list is then combined with a node list, as each I.D. of a variable in the edge list is connected to the corresponding I.D. in the node list.

Inputting the node and edge CSV files into VOSviewer allows the program to run its algorithm and construct a visual map of the network interactions based on the information extracted from the dataset. Additionally, the newly calculated edge and node files show each variable's two different 'weight' attributes. The two weight attributes are shown: (1) the 'Links' attribute and (2) the 'Total link strength’ (Van Eck & Waltman, Visualizing bibliometrics networks, 2014). In the case of this research, the 'Links' attribute is the number of collaboration links of a given variable with other individual variables. The 'Total link strength' attribute indicates the total strength of a variable's relationship with other variables. A variable with a higher value is regarded to have a stronger link, but the weaker links will always have the value of one.

The visualization software VOSviewer creates two different kinds of visualization maps: cluster and density maps. Each is chosen depending on the functionality and ease of understanding the data presented. The cluster maps are used to view network clusters, or groupings, though the density of the variables while displaying each variable separately; the more significant the ‘weight’ of a variable, the larger the label and node of the variable. Some of the lower ‘weight’ labels may not show by default, which is done by the program to avoid overcrowding labels on the map. The density map also shows the variables by their label, similar to the cluster map, with each variable visualized by a color that indicates the density of variables at that point. The colors range from blue to green to yellow to red, displaying the more significant number of variables in the neighborhood of a point and the higher the ‘weight’ of the neighboring variables, the closer the color is to red. The opposite is true, with the smaller the number of variables in the neighborhood of a point and the lower the ‘weights’ of the neighboring variables, the closer the color is to blue. In both maps, the relatedness of each variable is seen through their location on the map. The closer two variables are to each other, the stronger their connection, while the opposite displays the weaker connections.

VOSviewer maps can be constructed using a large dataset. A problem arises when the number of variables displayed is too many—making the image indistinguishable and challenging to understand, as seen in Figure 1. Due to this, the visual map in the findings that excludes any variables to focus on the centralized cluster will be the one displaying inter-organizational data. Figure 1 shows the complete version of the data, where even if the size is expanded, many variables are lost, and the trends to be observed are harder to discern. Nevertheless, the functionality of the maps created based on networks allows for the information on the Gulf of Mexico knowledge co-production to be presented in a relevant manner.

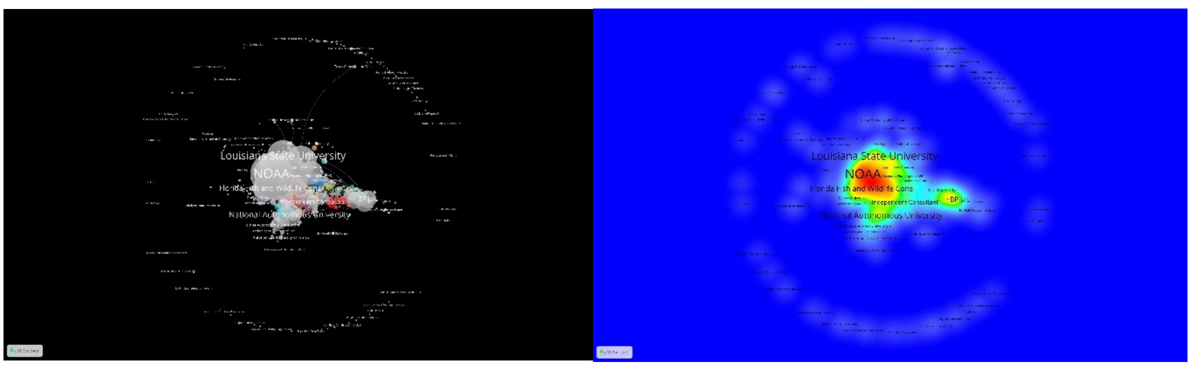

Figure 3: Sample of VOSviewer cluster and density map from inter-organizational data

Findings

Publication outputs, journals, and inter-organization collaboration

Published papers on the Gulf of Mexico were distributed to over 1,059 unique publication journals between 2000 and 2018. Table 1 lists the top 20 journals by the number of collaborative publications, with the top two being *Proceedings of the Annual Offshore Technology Conference* being the most prominent venue with 123 papers (2.88% of total output) and the *Proceedings - S.P.E. Annual Technical Conference and Exhibition* with 99 papers (2.32%). These two journal publishers not only cover engineering and energy production articles but are part of the conference proceedings. Two peer-reviewed journals follow these results, *Marine Ecology Progress Series* with 84 papers (1.97%) focused on all aspects of marine ecology and *PLOS One* with 79 papers (1.85%) focusing on any discipline within medicine and science.

Table 1: Top 20 journals publishing collaborative research

Journal Title	Papers (n)	Percent (%)
Proceedings of the Annual Offshore Technology Conference	123	2.88
Proceedings - S.P.E. Annual Technical Conference and Exhibition	99	2.32
Marine Ecology Progress Series	84	1.97
PLoS One	79	1.85
Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering – OMAE	69	1.62
Estuaries and Coasts	57	1.34
Continental Shelf Research	53	1.24
Journal of Geophysical Research: Oceans	48	1.13
Environmental Science and Technology	46	1.08
Deep-Sea Research Part II: Topical Studies in Oceanography	45	1.05
Journal of Coastal Research	45	1.05
Bulletin of Marine Science	41	0.96
Geophysical Research Letters	39	0.91
Estuarine Coastal and Shelf Science	38	0.89
Marine Biology	38	0.89
Marine Pollution Bulletin	38	0.89
SEG Technical Program Expanded Abstracts	37	0.87
Marine Chemistry	32	0.75
Marine and Coastal Fisheries	32	0.75
Marine and Petroleum Geology	31	0.73

The raw data from the SCOPUS database lists 1,648 publishing journals within the same years, with the top 20 journals in the publication of articles listed in Table 2. The top publisher still being the *Proceedings of the Annual Offshore Technology Conference* with 285 (3.57% of total output), showing a 162-paper difference. Investigation as to the changes in the output of the journal between the total amount of papers published to the number of collaborative papers resulted from the number of single-author articles, commonality in conference presentations, and articles compared to articles in peer-reviewed journals.

Table 2: Top 20 journals publishing research

Journal Title	Papers (n)	Percent (%)
Proceedings of the Annual Offshore Technology Conference	285	3.57
Oil and Gas Journal	188	2.35
Proceedings - S.P.E. Annual Technical Conference and Exhibition	188	2.35
Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering – OMAE	135	1.69
Marine Ecology Progress Series	117	1.46
SEG Technical Program Expanded Abstracts	115	1.44
PLoS One	95	1.19
Offshore Engineer	83	1.04
Continental Shelf Research	77	0.96
Journal of Coastal Research	77	0.96
Estuaries and Coasts	72	0.90
Estuarine, Coastal and Shelf Science	65	0.81
World Oil	65	0.81
Journal of Geophysical Research: Oceans	63	0.79
Environmental Science and Technology	62	0.78
Geophysical Research Letters	62	0.78
Bulletin of Marine Science	58	0.73
Deep-Sea Research Part II: Topical Studies in Oceanography	58	0.73
Marine Pollution Bulletin	55	0.69
SPE/IADC Drilling Conference, Proceedings	54	0.68

The trends in the distribution of annual publication outputs over our study period are presented in Figure 4. The total number of papers produced and the number of inter-organizationally co-coauthored papers show an increasing trend over time, consistent with other observations of research collaboration in other fields (Elango & Rajendran, 2012; Gorraiz & Schloegl, 2008; Song, Hickey, Temby, & Krantzberg, 2016; Lancho-Barrates & Cantu-Ortiz, 2019). A gap exists between publications involved in co-authored papers and the total number of papers published. This shows that while the increase in co-production is present, there are still other factors to consider in why this gap exists. Further investigation in locating what research field and journal publishers are absent or show low numbers of collaboration can illuminate where a need for co-production may benefit the field's innovation.

Figure 4: Total number of research papers and co-authored papers produced annually (2000-2018).

A total of 2,490 organizations were identified in all the collected articles. NOAA ranked the highest with 683 published papers (5.23% of all co-authored publications), connected to 431 different organizations with a total strength 1,325 number of links. It shows that NOAA is one of the critical organizations in knowledge co-production among the top 20 organizations listed in Table 3 that produced the most inter-organizationally authored research on the Gulf of Mexico but is also the most predominant organization displayed in Figure 5. The top seven organizations published well over 200 collaborative papers, consisting of two U.S. national government agencies, four U.S. universities, and one from Mexico. Closely followed by a mix of educational institutions, private corporations, and national and state government agencies. The number of inter-organizational network links in these organizations follows the paper publishing trend, but this is not consistently seen in the case of *Schlumberger Ltd*, which has 108 published papers but is only connected to 74 different organizations and has a total strength of 156 number of links. The organization label is not seen in Figure 5 as other organizations overshadow it. The density map created by the inter-organizational network displays two prominent clusters, the larger of the two on the left side of the map, with NOAA centralized to various types of organizations. It shows a deeper red color, indicating the higher level of co-production occurring amongst those in the cluster. In contrast, the smaller cluster on the right is predominantly private organizations with a fair amount of co-production, seeing that the coloration is yellow and not blue. The fact that there is a distinction between these two displays a lack of connections among the organizations making up these clusters.

Table 3: Top 20 organizations

Organization	Papers (n)	Percent (%)	Link	Total link strength
National Oceanic and Atmospheric Administration (NOAA)	683	5.23	431	1325
Louisiana State University	333	2.55	263	751
Texas A&M University	307	2.35	320	738
US Geological Survey	305	2.33	257	613
University of Southern Mississippi	213	1.63	187	512
National Autonomous University of Mexico (UNAM)	211	1.62	206	373
University of South Florida	203	1.55	202	535
University of Texas at Austin	149	1.14	223	414
Florida State University	136	1.04	214	418
B.P. plc	133	1.02	129	233
University of Miami	133	1.02	158	390
University of South Alabama	131	1.00	128	281
Chevron Corp	128	0.98	110	200
Shell Corporation	125	0.96	112	182
University of North Carolina	120	0.92	195	414
University of Georgia	115	0.88	161	379

Table 3, cont.

Schlumberger Ltd	108	0.83	74	156
Florida Fish and Wildlife Conservation Commission	106	0.81	129	307
US Environmental Protection Agency	102	0.78	111	184
University of Florida	100	0.77	150	296

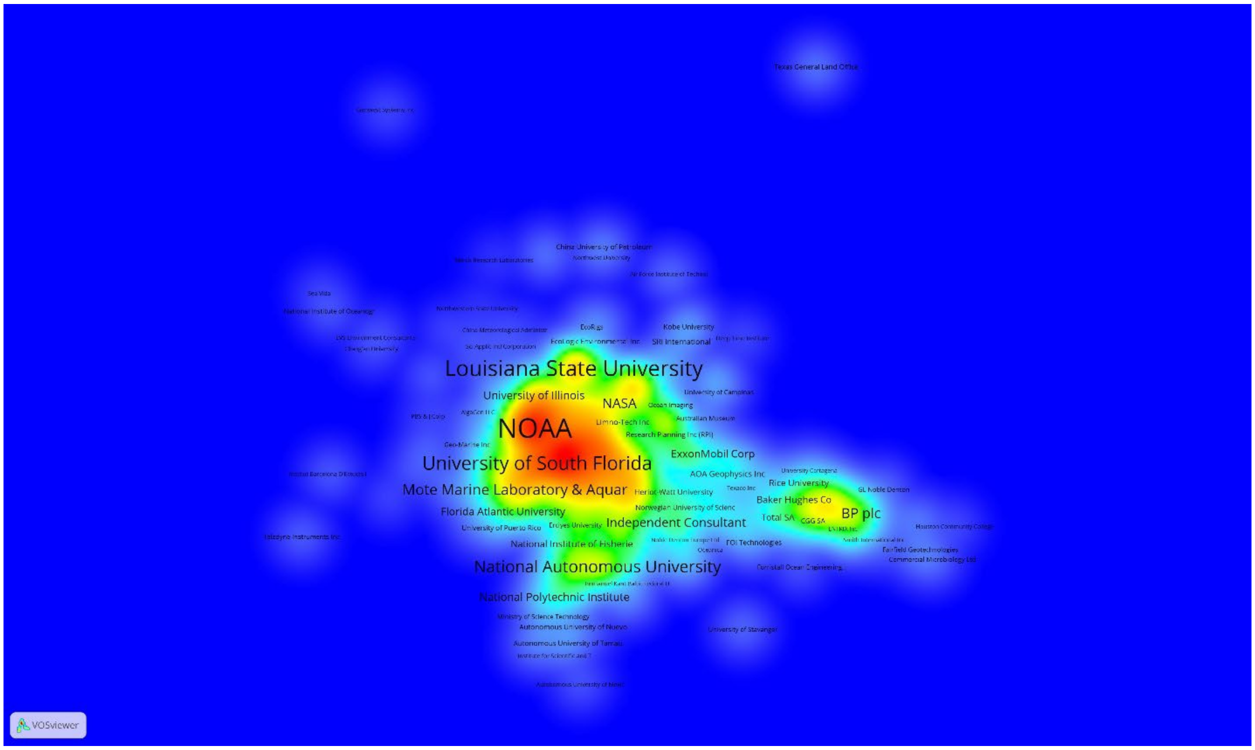

Figure 5: Inter-organizational research collaboration density map

In the interest of observing the output of research papers numbers from organizations located within the Gulf of Mexico region and the rate of co-production occurring only within those organizations, Table 4 lists the top 20 resulting organizations. It is found that many of the same top 20 organizations from Table 3 are present but have a difference in the number of network links within organizations of the region. Figure 6 presents the resulting network of the inter-organizational co-production in the form of density visualization, forming three distinct clusters. The redder the color, the greater the level of collaboration, with the closer distance showing the strength of their connections. In other words, the closer the organizations are to each other on the map, the greater the number of times those organizations have worked with one another. The largest cluster on the left shows a light level of co-productions within various institutions and NOAA with a deep red coloration. While in contrast, the second largest cluster on the far right is notably not as intense but is darker in the yellow coloration than the smaller central cluster. An interesting note is a cluster in the middle, with the *University of Texas at Austin* being the most prominent organization, which displays through its location the probability of the organizations forming this cluster being facilitators between the other two distinctly larger clusters.

Table 4: Top 20 organizations located in the Gulf of Mexico region

Organization	Papers (n)	Percent (%)	Link	Total link strength
National Oceanic and Atmospheric Administration (NOAA)	683	10.3	136	685
Louisiana State University	333	5.02	99	408
Texas A&M University	307	4.63	116	366
US Geological Survey	305	4.6	99	332
University of Southern Mississippi	213	3.21	69	285
University of South Florida	203	3.06	63	312
University of Texas at Austin	149	2.25	88	208
Florida State University	136	2.05	71	182
University of Miami	133	2.01	64	243
University of South Alabama	131	1.98	50	156
Florida Fish and Wildlife Conservation Commission	106	1.6	52	200
US Environmental Protection Agency	102	1.54	49	100
University of Florida	100	1.51	61	176
Texas A&M University at Corpus Christi	99	1.49	52	147
Schlumberger Ltd	96	1.45	45	85
U.S. Naval Research Laboratory	87	1.31	34	132
Shell Corporation	86	1.3	47	92
U.S. Department of Agriculture - A.R.S.	80	1.21	22	56
B.P. plc	79	1.19	49	81
Louisiana Universities Marine Consortium (LUMCON)	79	1.19	38	129

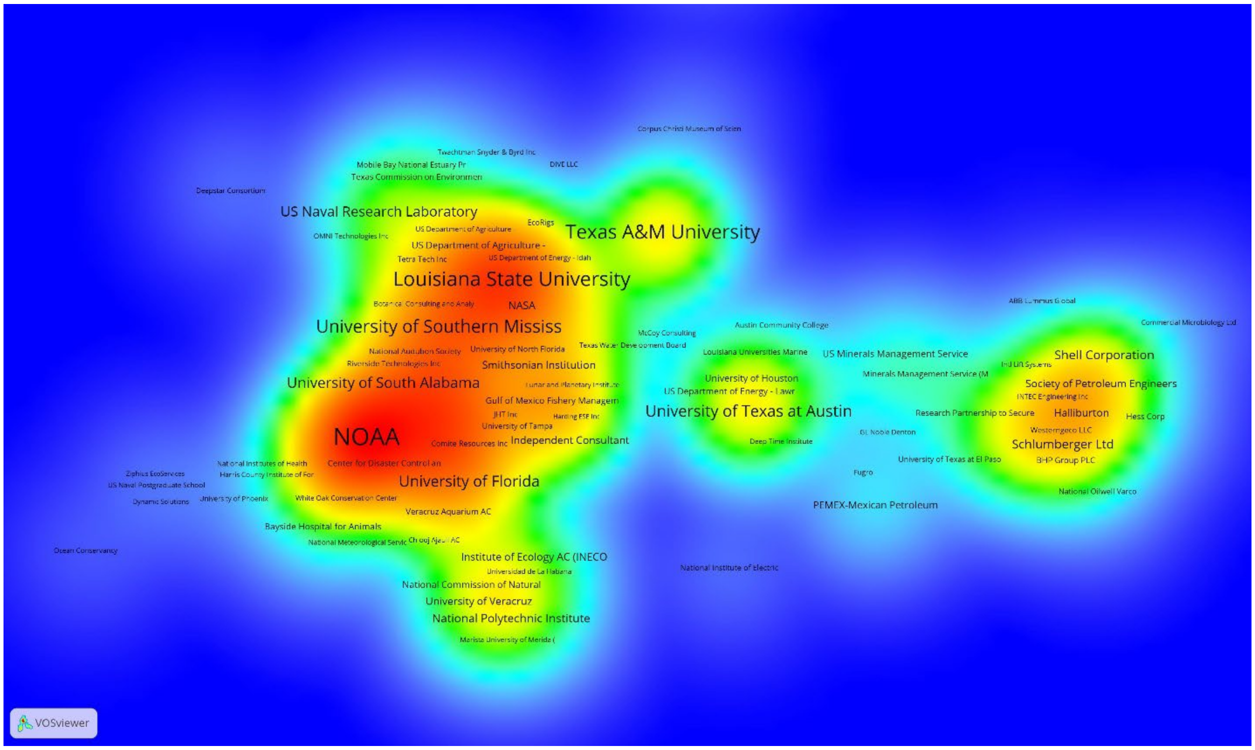

Figure 6: Inter-organizational research collaboration in the Gulf of Mexico region density map

Geographical and Inter-jurisdictional collaboration

Using STATA to calculate the frequency for the identified 82 countries, 52 states, and 21 jurisdictions allow us to see the number of authors and organizations involved in research co-production and the level of collaboration. Table 5 shows in the first three rows the countries located in direct contact with the Gulf of Mexico region, with the United States having the most considerable amount of authorship (74.19% of all authors), organizations (53.25%) with 75 different collaboration links and Mexico is the second largest with 7.04% authorship, 4.62% of organizations and 40 different links. On the opposite end is Cuba, which, if not for being a focus area, would not have been part of the top 10 countries listed. It has only 17 authors (0.13% of all authors) from nine organizations (0.36%) and has collaborative links with only three countries.

The United Kingdom, Canada, and the following countries in Table 5 show the top 10 authorship involved in the co-production.

The density map visualizing the country co-production network in Figure 7 shows that the United States has high levels of research collaboration, with its node being a dark red color. Additionally, combined with the observations in Table 5, its position in the midway area of all the nodes displayed in the map shows the potential role as a facilitator for science research occurring in the Gulf of Mexico region. It is the only node with weak links, accounting for only one or two collaborations with each country listed on the far right. Explaining the large gap between the United States has a stronger connection with the countries in the largest cluster of countries, such as Mexico, Canada, Germany, and the United Kingdom. The countries of this large cluster, while not producing as much in comparison to the U.S., their distance show stronger relations with one another.

An interesting thing to note in these findings is that the number of times the United States and Mexico collaborated is 460, a mere 12% of all U.S. collaboration links. At the same time, both have only collaborated with Cuba 19 and 18 times, respectively. Adding this information to Figure 7, we see a disparity in co-production among these three countries. The strength of collaboration between the node representing the United States is weaker as its located further away from Cuba and Mexico than Canada.

Table 5: Gulf of Mexico region countries and top 10 countries in geographical collaboration

Country	Author (n)	Percent (%)	Organization (n)	Percent (%)	Link (n)	Total link strength (n)
United States	9,692	74.19	1,326	53.25	75	3722
Mexico	920	7.04	115	4.62	40	767
Cuba	17	0.13	9	0.36	3	39
United Kingdom	331	2.53	116	4.66	38	731

Table 5, cont.

Canada	299	2.29	81	3.25	32	740
China	205	1.57	72	2.89	24	389
Germany	188	1.44	66	2.65	29	545
France	158	1.21	71	2.85	32	369
Brazil	121	0.93	44	1.77	20	151
Netherlands	119	0.91	36	1.45	22	318
Norway	114	0.87	44	1.77	18	165
Spain	97	0.74	60	2.41	31	287
Australia	96	0.73	41	1.65	26	280
Italy	81	0.62	42	1.69	29	280
Japan	73	0.56	39	1.57	17	263
South Korea	52	0.40	29	1.16	12	87
Switzerland	40	0.31	15	0.06	13	128
Taiwan	35	0.27	11	0.44	13	92
Malaysia	25	0.19	9	0.36	11	28
Colombia	24	0.18	17	0.68	8	26
Russia	24	0.18	13	0.52	10	24

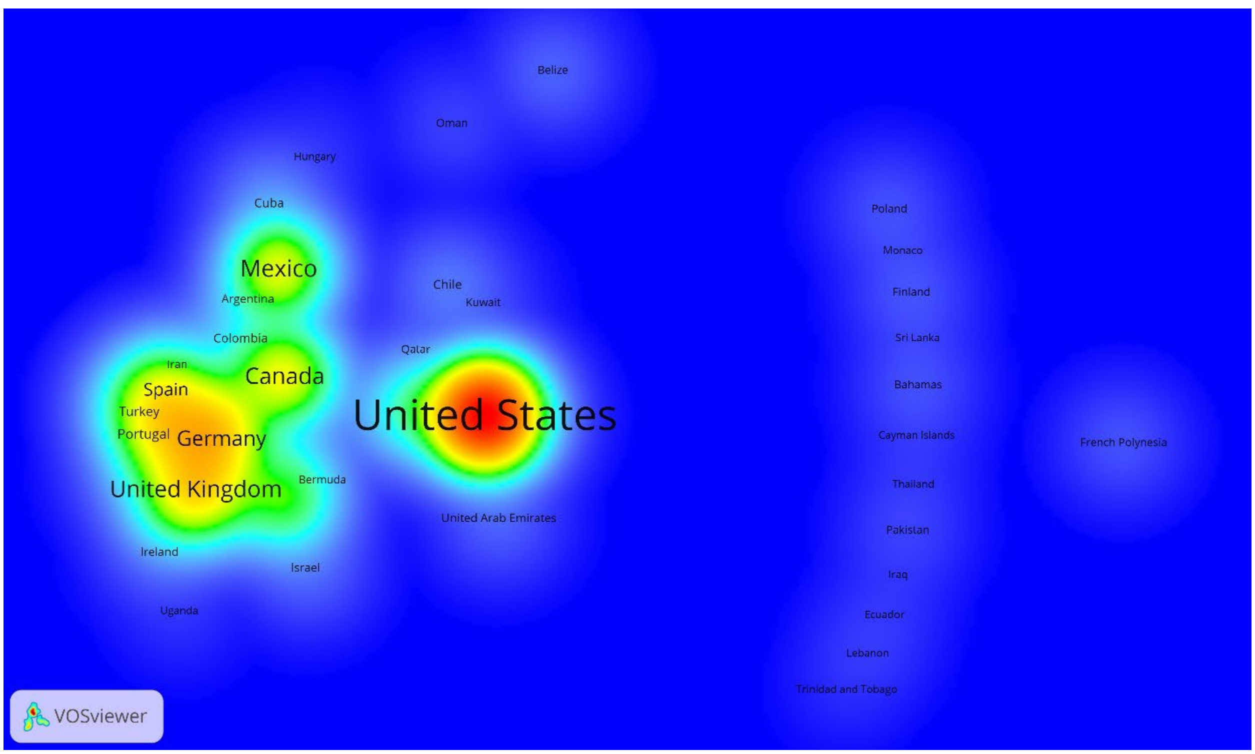

Figure 7: Country geographical collaboration density map

Geographically narrowing down to the three countries which border the Gulf of Mexico, a total of 52 states were identified between the United States and Mexico. Table 6 shows the top 10 states with the most significant number in authorship from each respective country, in addition to the data from Cuba. Many of the states listed are those geographically considered within the research region. However, for Mexico, the top state listed for authorship (277), number of organizations (26), and collaboration links (77) is the capital of Mexico City. A state geographically located in the center of the country and not considered a coastal state. This state's prominence can be explained by looking back at the results from Table 3, where the internationally recognized *National Autonomous University of Mexico (UNAM)* is listed in the Top 20 co-production organization.

Meanwhile, the country of Cuba holds the lowest number of authorship (17), organizations (9), and one of the two lowest numbers of collaboration links (14) listed in the table. In combination with Table 6 and Figure 8, the two states with the highest degree of collaboration, Texas and Florida, are not as close in strength as Louisiana and Mississippi are to Florida. However, the location of Florida on the overall map reflects a degree of the state's co-production centralization connecting the states on either side.

In addition, Veracruz is the second top Mexican state with 132 authorships, 9 organizations, and 61 collaboration links; it is located in a cluster of Mexican states on the far left of Figure 8. An apparent distance between this cluster and those that make up the U.S. states further shows the weakness between these two countries' co-production.

Table 6: Top 10 United States and Mexico states and Cuba

Country	Jurisdiction	Author (n)	Organization (n)	Link (n)	Total link strength (n)
United States	Texas	1,991	306	108	2606
	Florida	1,496	111	104	2734
	Louisiana	810	65	87	1413
	California	631	102	99	1569
	Mississippi	567	27	71	1065
	North Carolina	385	32	79	1044
	Massachusetts	300	50	77	804
	Maryland	291	40	60	626
	Alabama	272	21	63	551
	Colorado	261	47	71	638
Mexico	Mexico City	277	26	77	804
	Veracruz	132	9	61	244
	Yucatan	122	9	53	202
	Campeche	75	6	31	105
	Baja California	41	5	32	90
	Baja California Sur	39	2	34	84
	Quintana Roo	35	9	14	27
	Tabasco	31	3	31	128
	Morelos	29	3	19	39
	Tamaulipas	23	5	16	29
Cuba	Cuba	17	9	14	40

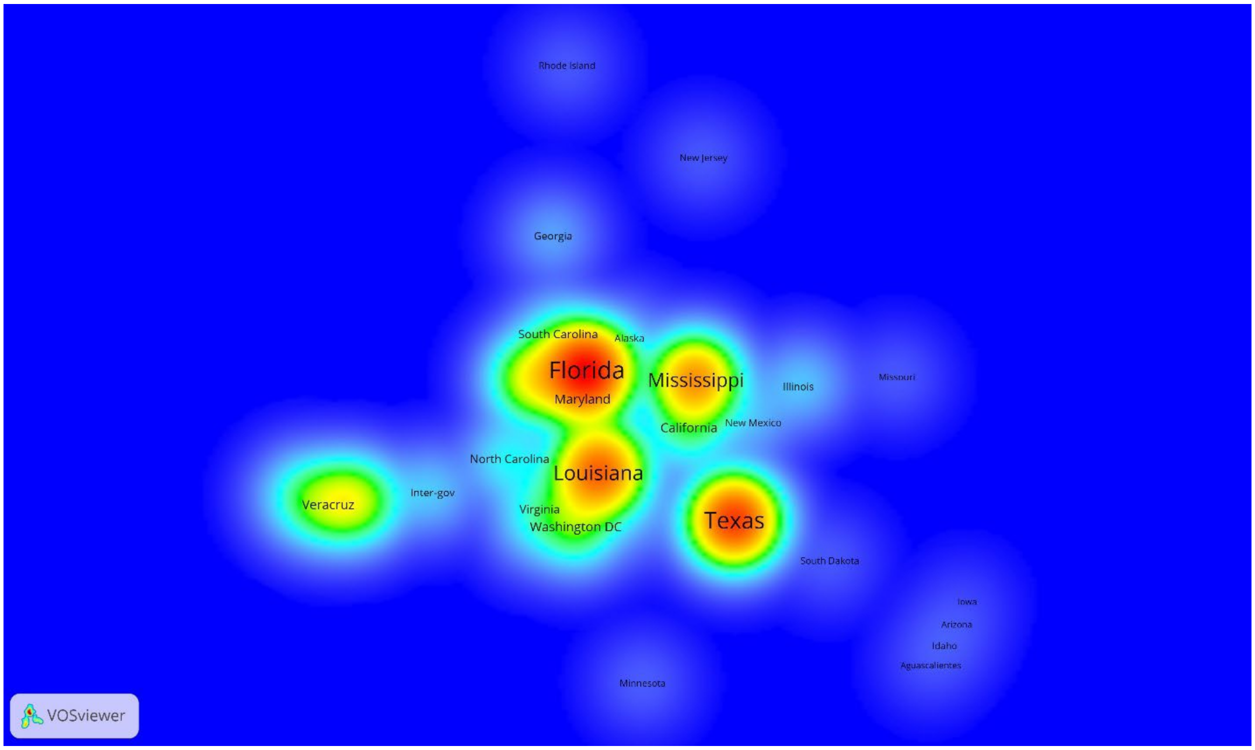

Figure 8: U.S. and Mexican state collaboration density map

The findings on the 16 jurisdictions that make up the Gulf of Mexico region are seen in Table 7. The network of inter-jurisdictional collaboration is visualized through the density and cluster maps of Figure 9 and Figure 10, respectively. As displayed by the node color, the results highlight a high degree of research collaboration, making a cluster made of the U.S. national level with regional collaborative authorship of 1,784 within 61 organizations, Alabama with 245 of 18 organizations, and Louisiana with 672 of 59 organizations nodes. The next closest cluster is made by the state of Texas, with 1,889 regional authorship collaborations of 304 organizations, and Florida with 1,152 of 106. The short distance between clusters made of the described U.S. jurisdictions shows close ties in their co-production, with only the U.S. state of Mississippi having the weaker connection.

Notably apparent is the clustering by country, further exhibiting the fragmentation between Mexico and U.S. co-production. On the other hand, the strength of collaboration between inter-governmental organizations, Mexican states, and Cuba shows collaborative robustness that is not as strong in the U.S. At the same time, their regional authorship is lower and collaborative production is lower, as shown by the coloration of the cluster.

In the observation of Cuba and its connection, the stronger ties lie with the Mexican states with 29 total links; there are still connections with Florida and Texas. Continuing that observation, the intermediate location of Texas and Florida in the map between all the clusters reflects their role as facilitators of transboundary innovation and production of scientific knowledge

Table 7: Gulf of Mexico inter-jurisdictional collaboration

Country	Jurisdiction	Author (n)	Organization (n)	Link (n)	Total link strength (n)
United States	Alabama	245	18	11	475
	Florida	1,152	106	17	2521
	Louisiana	672	59	14	1300
	Mississippi	358	24	11	853
	Texas	1,889	304	15	2545
	US Federal	1,784	61	16	3423
Mexico	Campeche	58	4	12	88
	Quintana Roo	29	8	14	58
	Tabasco	26	3	12	34
	Tamaulipas	18	5	8	24
	Veracruz	126	7	15	224
	Yucatan	119	9	14	210
	MX National	115	18	18	218
Cuba	Cuba	11	5	10	28
	Cuba National	6	4	9	18
Other	Inter-gov	21	17	10	90

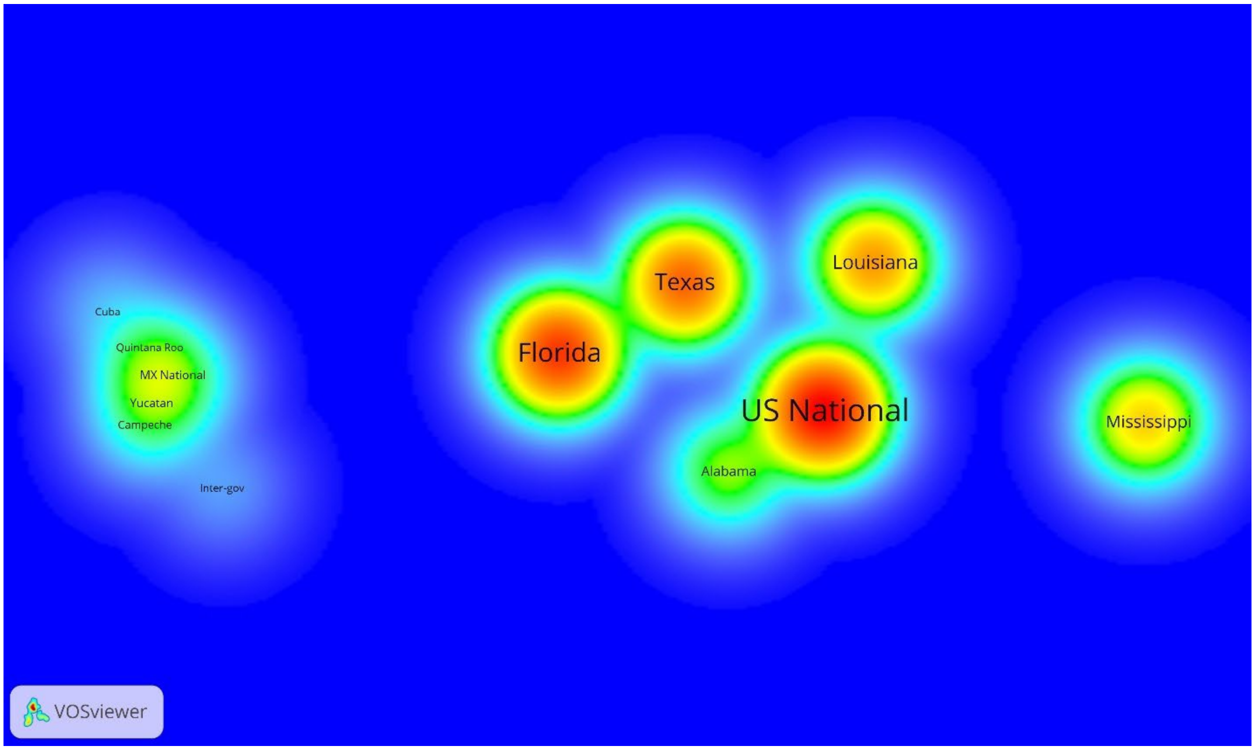

Figure 9: Inter-jurisdictional Gulf of Mexico region density map

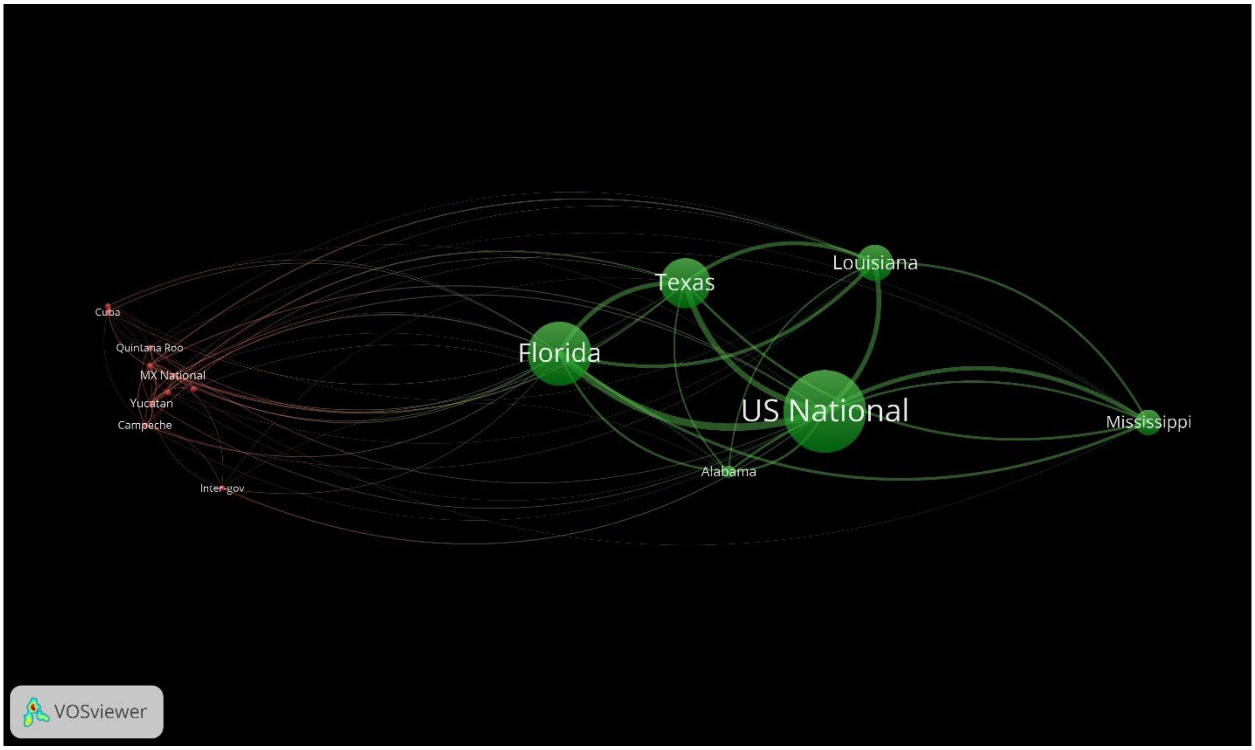

Figure 10: Inter-jurisdictional Gulf of Mexico region collaboration cluster map

Inter-institutional collaboration

The basic overview of inter-institutional collaborations occurring in Gulf of Mexico scientific research shows that the largest comes from educational institutions, which make up 44.32% of the overall authorship with 1,052 identified organizations, followed by both federal and private institutions, 28.75% with 209 and 19.25% with 924 organizations respectively. As an initial observation, the interest of the analysis in institutions will be further focused on the institutions located in the Gulf of Mexico region to measure their degree of collaboration.

Table 8: Inter-institutional collaboration

Institution	Author (n)	Percent (%)	Organization (n)	Percent (%)	Total link strength (n)
Education	2,938	44.32	1052	42.25	6769
National	1,906	28.75	209	8.39	4500
Inter-gov	21	0.32	17	0.68	90
NPO	225	3.39	192	7.71	1729
Private	1,276	19.25	924	37.11	2744
Sub-National	263	3.97	96	3.86	1080

The inter-institutional collaboration network diagrams are visualized in Figure 11 as a cluster map and a density map in Figure 12. Measuring through co-authorship and the number of organizations in Table 9, a substantial degree of co-production occurs through the 405 different U.S. Educational institutions, the 571 International Education institutions, and 61 U.S. National institutions. These levels of collaboration are followed in both the degree of coloration and node size by the 646 U.S. private institutions. Comparatively, weak levels of collaboration can be seen in other institutions, particularly Cuba's single private institution with only two different links of collaboration and the 10 international sub-national institutions with nine different links. As seen in the geographical and inter-jurisdictional density maps, there is a clear distinction between the countries, supporting the suggestion of research fragmentation occurring between the countries.

Surprisingly, the connections between federal and academic research in the United States show it not being as strong in collaboration compared to their Mexican equivalents based on Figure 11. Nevertheless, by observing Figure 12, a strong connection is displayed by the thickness of the link connecting the two nodes. It makes us believe that while the distance displayed in the density map can show the strength of a relation between two nodes, there is a balance to be made by accounting for the total number of different links each has. In this case, the node for U.S. Education is placed in a place that shows its higher part of a centralized network.

Table 9: Gulf of Mexico inter-institutional collaboration

Institution	Author (n)	Percent (%)	Organization (n)	Percent (%)	Link (n)	Total link strength (n)
U.S. Education	5,183	39.67	405	16.27	17	6734
US Private	1,934	14.80	646	25.94	14	2551
US National	1,785	13.66	61	2.45	16	3422
US NPO	419	3.21	131	5.26	15	1262
US Sub-National	368	2.82	83	3.33	13	1014
MX Education	748	5.73	72	2.89	19	711
MX National	115	0.88	18	0.72	17	218
MX Private	35	0.27	10	0.4	9	98
MX N.P.O.	17	0.13	10	0.4	10	47
MX Sub-National	3	0.02	3	0.12	6	9
C.U.B. Education	9	0.07	3	0.12	6	22
C.U.B. National	6	0.05	4	0.16	7	18
CUB NPO	1	0.01	1	0.04	3	3
C.U.B. Private	1	0.01	1	0.04	2	2
INT Education	1,387	10.62	571	22.93	15	3082
INT Private	599	4.59	267	10.72	14	838
INT National	296	2.27	127	5.1	14	1013
INT NPO	119	0.91	50	2.01	14	443
INT Sub-National	18	0.14	10	0.4	9	59
Inter-gov	21	0.16	17	0.68	12	90

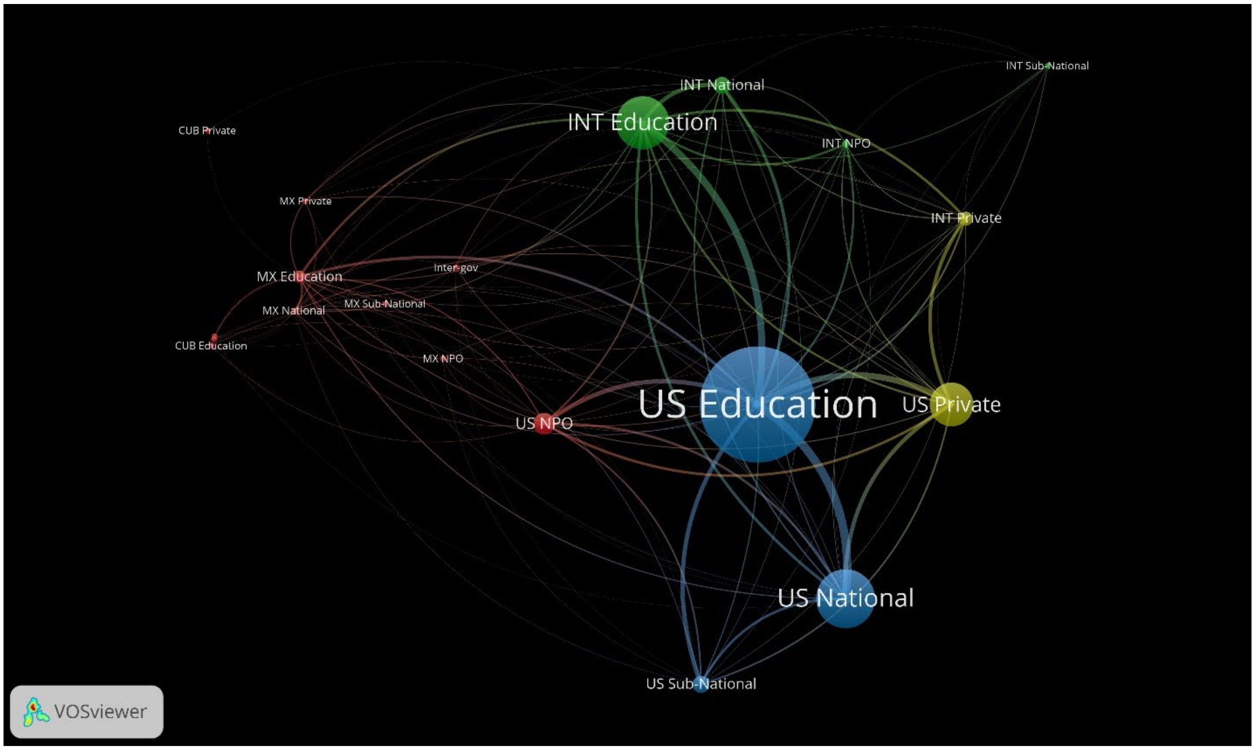

Figure 11: Inter-institutional collaboration cluster map

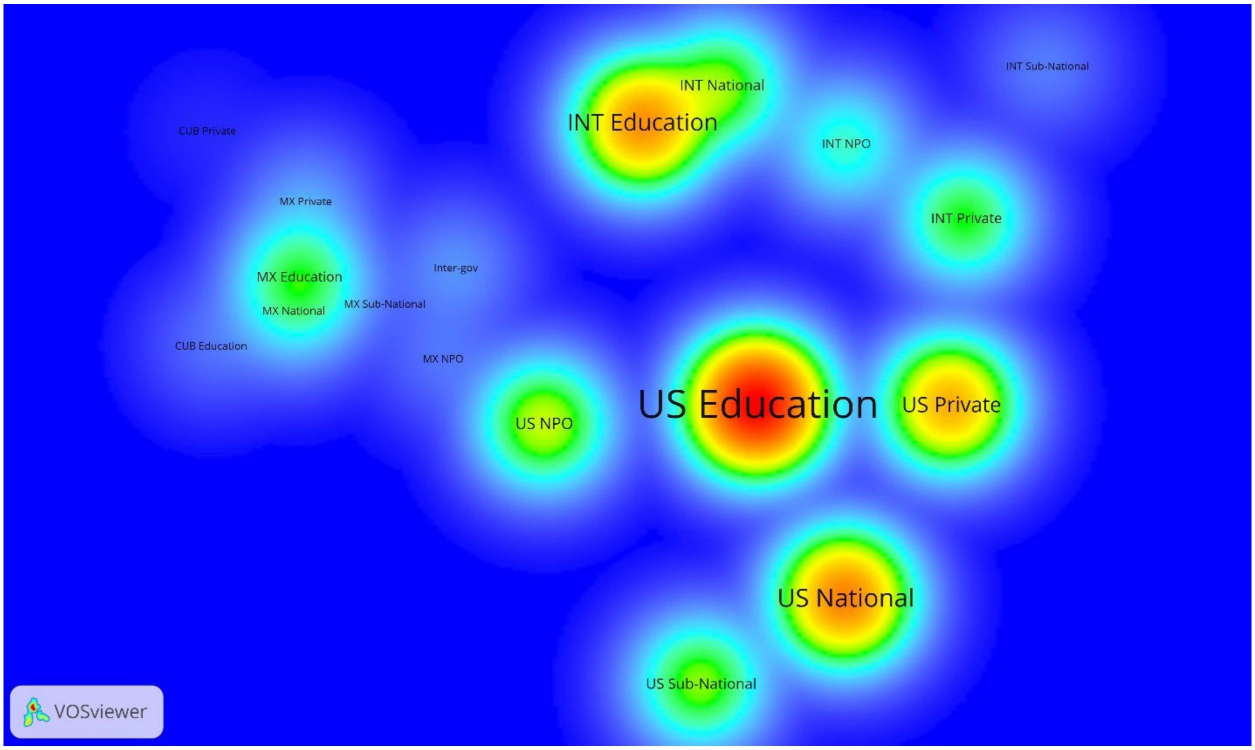

Figure 12: Inter-institutional collaboration density map

Limitations

Bibliometric analysis is an effective tool to obtain a primary analysis of the networks and knowledge dissemination occurring in the Gulf of Mexico, but by no means is a replacement for qualitative evaluation methods. Some limitations of using this method relate to how research and scholarly publishing works are presented. The data published from scientific databases such as SCOPUS can contain errors affecting the data if not managed properly (Donthu, Kumar, Mukherjee, Pandey, & Lim, 2021). To ascertain that these errors are mitigated, careful work cleaning the data must be done before any analysis, which depending on the dataset size, can be time-consuming.

Secondly, the quantitative nature of bibliometric analysis can introduce unknown limits, as different areas of study have different publication frequencies and citation cultures. While scholars recognize SCOPUS and WoS as established databases, there are still publications that get missed. An example is the publications available in the repository library by NOAA, where some of the articles available are published in recognizable journals, while the documents labeled Technical Memorandums do not. The need for further analysis is always recommended as different variables can cause sources to be unaccounted for, such as the simple limitation due to language barriers or even the article source.

CHAPTER IV

DISCUSSION AND CONCLUSION

In the work to improve transboundary networks of the Gulf of Mexico, results reveal that the production of scientific research is steadily continued to rise over the years and has a steady network within each country. However, a shortcoming in the co-production of scientific research between the regional countries was discovered.

As part of the results, the volume of papers related to issues in the Gulf of Mexico that researchers co-authored has increased over the last 18 years, with most collaborations being done between federal, private, and educational institutions. The extent of fragmentation in scientific co-production between Mexican and U.S. institutions draws a surprise. One would have expected that with the long history of binational policy efforts over water and border resources, the same intensity would be placed on the Gulf of Mexico system (Correa-Cabrera & Konrad, 2020; Coronado & Mumme, 2020). These results suggest a need to increase the attention towards establishing cross-border scientific relations for the co-production, which would facilitate opportunities to consolidate research synthesis and increase the dissemination of scientific knowledge along relevant stakeholders. Nevertheless, it is understood that reports on the degradation between the United States and Mexico relations in recent years have not been lacking, not even touching on the US-Cuba issues. There is still a case to be made that many stakeholders still see the importance of cross-border initiatives and environmental cooperation between the countries (Temby & Stoett, 2017). A recognized cross-border collaboration partnership over the Gulf of Mexico is currently lacking, it is proven that co-production efforts between U.S. and Mexico are practical. An example of such a transboundary program is the International Boundary and Water Commission. While the commission has been considered overly centralized and bureaucratic, its mission follows the responsibilities stated in the Water Treaty of 1944. The efforts made following this mission have shown progress through reaching out to various stakeholders and advancing water management applications (Pineda Pablos, Mumme, Rivera Torres, Vega, & Hernandez, 2020; Mumme, 2003). Creating opportunities to co-produce scientific knowledge is imperative in addressing many issues in managing a region. These kinds of opportunities help in the disclosure and discussion of research projects and allow a better understanding of the trade-offs in decision-making.

An interesting aspect in observations found for the country of Cuba, whose fragmentation with the United States was expected, shows that the data for co-production with Mexico is robust in both jurisdictional and institutional systems. The finding creates an opportunity for stakeholders in the U.S. to strengthen its link with Cuba through Mexico. More in-depth analysis of Cuba, its production of scientific knowledge, and current collaborative programs would need to be done.

The likely challenges in the co-production efforts tend to fall on the stakeholders' political and social will, which can make misguided recommendations in management practices when lacking binational unity due to lacking a sturdy scientific foundation and unbiased objectivity. Given the importance of good cooperation in management, it remains imperative that common objectives are adopted when developing and implementing cooperative programs. The question remains as to why scientific collaboration at the transnational level appears limited in the Gulf of Mexico region and how one can facilitate a more inclusive and integrated system where the enhancement of innovation will support a more sustainable region of governance.

Even when using the bibliometric analysis method and the size of the dataset, this research is only a preliminary analysis of the networks and knowledge dissemination occurring in the Gulf of Mexico. Additionally, it is only a partial indicator as further analysis is required to calculate the level of trust among stakeholders involved in cross-national governance. The following step to do with the data obtained in this research is to look further into the networks already present and investigate the state of quality of the collaborations, much like was done with the research in the Great Lakes (Song, Temby, Dongkyu, Savendra Cisneros, & Hickey, 2019; Sohns, Hickey, de Varies, & Temby, 2021). In disseminating research, we would need to broaden our scope to understand better the different actors in producing scientific knowledge.

REFERENCES

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981-1999. *Research Policy, 34*(3), 259-285.

Alvarez Torrez, P., Rabalais, N. N., Pina Gutierrez, J. M., & Pardon Lopez, R. M. (2017). Research and community of practice of the Gulf of Mexico large ecosystem. *Environmental Development, 22*, 166-174.

Banco de Mexico. (2018). *Reporte sobre las economias regionales.* Retrieved from Banco de Mexico: Publicaciones: https://www.banxico.org.mx/publicaciones-y-prensa/reportes-sobre-las-economias-regionales/%7B22964087-C9F0-6A32-4ACC-EF2A104EB59D%7D.pdf

Beier, P., Hansen, L. J., Helbrecht, L., & Behar, D. (2017). A how-to guide for coproduction of actionable science. *Conservation Letters: A journal of the Society for Conservation Biology, 10*(3), 288-296.

Belter, C. W. (2013). A bibliometric analysis of NOAA's Office of Ocean Exploration and Research. *Scientometrics, 95*, 629-644.

Bennett, L. M., & Gadlin, H. (2012). Collaboration and Team Science: From theory to practice. *Journal of Investigative Medicine, 60*(5), 768-775.

Bornmann, L., & Ozimek, A. (2012). STATA commands for importing bibliometric data and processing author address information. *Journal of Informetrics, 6*(4), 505-512.

Cavadas, A. (2020). Visualizing the collaboration network of a European marine research infrastructure. *University of Porto Journal of Engineering, 6*(2), 98-118.

Cinner, J. E., & Barnes, M. L. (2019). Social dimensions of resilience in social-ecological systems. *One Earth, 1*(1), 51-56.

Conley, A., & Moote, M. (2003). Evaluating Collaborative Natural Resource Management. *Society & Natural Resources*, 371-386.

Coronado, I., & Mumme, S. (2020). Environmental governance at the US-Mexico Border. In T. Payan, & P. L. Cruz, *Binational Commons: Institutional Developments and Governance on the U.S.-Mexico Border* (pp. 159-190). The University of Arizona Press.

Correa-Cabrera, G., & Konrad, V. (2020). *North American borders in comparative perspective.* University of Arizona Press.

Cruz, I., & McLaughlin, R. J. (2008). Contrasting marine policies in the United States, Mexico, Cuba and the European Union: Searching for an integrated strategy for the Gulf of Mexico region. *Ocean & Coastal Management, 51*(12), 826-838.

Delacamara, G., O'Higgins, T. G., Lago, M., & Langhans, S. (2020). Ecosystem-based Management: Moving from Concept to Practice. In T. O'Higgins, M. Lago, & T. DeWitt, *Ecosystem-based managment, Ecosystem Services and Aquatic Biodiversity* (pp. 39-60). Springer.

Djenontin, I., & Meadow, A. M. (2018). The art of co-production of knowledge in environmental sciences and managment: lessons from international practice. *Environmental Managment, 61*, 885-903.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285-296.

Elango, B., & Rajendran, P. (2012). Authorship trends and collaboration pattern in the marine sciences literature: A scientometric study. *International Journal of Information Dissemination and Technology, 2*(3), 166-169.

Glänzel, W., & Schubert, A. (2004). Analysing Scientific Networks Through Co-Authorship. *Handbook of Quantitative Science and Technology Research*, 257-276.

GoMA, Gulf of Mexico Alliance. (2019, April 22). *Gulf of Mexico Alliance*. Retrieved from Gulf of Mexico Alliance: http://gulfofmexicoalliance.org/

Gorraiz, J., & Schloegl, C. (2008). A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science. *Journal of Information Science*, 715-725.

Gunton, T., Peter , T., & Day, J. (2007). Evaluating Collaborative Planning: A Case Study of a Land and Resource Managment Planning Process. *Environments Journal, 34*(3), 19-37.

Harzing, A. W., & Alkangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics, 106*, 787-804.

Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library & Information Science Research*, 323-342.

Kalhor, G., Sarijalou, A. A., Sadr, N. S., & Bahrak, B. (2022). A new insight to the analysis of co-authorship in Google Scholar. *Applied Network Science, 7*(21).

Koseoglu, M. A. (2016). Growth and structure of authorship and co-authorship network in the strategic management realm: Evidence from the Strategic Management Journal. *BRQ Business Research Quarterly, 19*(3), 153-170.

Lancho-Barrates, B. S., & Cantu-Ortiz, F. J. (2019). Science in Mexico: a bibliometric analysis. *Scientometrics, 118*, 499-517.

Lubell, M. (2014). Collaborative partnerships in complex institutional systems. *Environmental Sustainability*, 41-47.

Mihiri Shashikala, H. B., George, R., & Shujaee, K. A. (2015). Outlier detection in network data using the Betweenness Centrality. *SoutheastCon 2015.* Fort Lauderdale: IEEE. doi:10.1109/SECON.2015.7133008

Milesi, C., Brown, K. L., & Schneider, B. L. (2016). Collecting, processing, and analyzing bilbiometric data for program evaluation. *SAGE Research Methods Cases Part 2*.

Miranda-Gonzalez, A., Aref, S., Theile, T., & Zagheni, E. (2020). Scholarly migration within Mexico: analyzing internal migration among researchers using Scopus longitudinal bibliometric data. *EPJ Data Science, 9*(34).

Mongeon, P., & Paul-Hus, A. (2015). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics, 106*(1), 213-228.

Mumme, S. P. (2003). Revising the 1944 Water Treaty: Reflections on the Rio Grande drought crises and other matters. *Journal of the Southwest, 45*(4), 649-670.

Narin, F., & Hamilton, K. S. (1996). Biliometric performance measures. *Scientometrics, 36*, 293-310.

Newman, M. E. (2001). The structure of scientific collaboration networks. *Applied Mathematics, 98*(2), 404-409.

Nostrom, A. V., Cvitanovic, C., Lof, M. F., West, S., Wyborn, C., Balvanera, P., . . . Osterblom, H. (2020). Principles for knowledge co-production in sustainability research. *Nature Sustainability, 3*, 182-190.

Noyons, E. C., Moed, H. F., & Van Raan, A. F. (1999). Integrating research performance analysis and science mapping. *Scientometrics, 46*, 591-604.

Pineda Pablos, N., Mumme, S., Rivera Torres, M., Vega, M. V., & Hernandez, O. (2020). Professional Expertise in the Management of International Freshwater Commissions: The Case of the. *Environmental Science and Policy, 114*, 43-51.

Powell, K. R., & Peterson, S. R. (2017). Coverage and quality: A comparison of Web of Science and Scopus databases for reporting faculty nursing publication metrics. *Nursing Outlook, 65*(5), 572-578.

Pranckute, R. (2021). Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World. *Publications (Basel), 9*(12), 12.

Ramos-Rodriguez, A.-R., & Ruiz-Navarro, J. (2004). Changes in the intellectual structure of strategic management research: a bibliometric study of the Strategic Management Journal, 1980–2000. *Srategic Management Journal, 25*(10), 981-1004.

Rogers, G., Szomszor, M., & Adams, J. (2020). Sample size in bibliometric analysis. *Scientometrics, 125*, 777-794.

Sacramento River Science Partnership. (2021). Retrieved from Sacramento River Science Partnership: https://www.sacriverscience.org/

Sherman, K. (2014). Toward Ecosystem-based Managment (EBM) of the World's Large Marine Ecosystem During Climate Change. *Environmental Developtment*, 43-66.

Sohns, A., Hickey, G. M., de Varies, J. R., & Temby, O. (2021). Methodological challenges in studying trust in natural resources management. *Land, 10*(12).

Song, A. M., Hickey, G. M., Temby, O., & Krantzberg, G. (2016, February). Assessing transboundary scientific collaboration in the Great Lakes of North America. *Journal of Great Lakes Research, 42*(1), 156-161.

Song, A. M., Temby, O., Dongkyu, K., Savendra Cisneros, A., & Hickey, G. M. (2019). Measuring, Mapping and Quantifying the Effects of Trust and Informal Communication on Transboundary Collaboration in the Great Lakes FIsheries Policy Network. *Global Envrionmental Change*, 6-18.

Su, M., Peng, H., & Li, S. (2021). A visualized bibliometric analysis of mapping research trends of machine learning in engineering (MLE). *Expert Systems with Applications, 186*.

Temby, O., & Stoett, P. (2017). *Towards continental environmental policy? North American Transnational Networks and Governance.* New York: State University of New York Press.

UCLA: Statistical Consulting Group. (2021). *How can I create all pairs within groups*. Retrieved from UCLA: Statistical Methods and Data Analytics: https://stats.oarc.ucla.edu/stata/faq/how-can-i-create-all-pairs-within-groups/

United States. National Centers for Coastal Ocean Science. (2022). *NOAA RESTORE Science Program*. Retrieved from National Oceanic and Atmospheric Administration: https://restoreactscienceprogram.noaa.gov/

United States. National Marine Fisheries Service. (2020). *2020 Fisheries of the United States.* Retrieved from National Oceanic and Atmospheric Administration: Institutional Repository: https://repository.library.noaa.gov/view/noaa/40953

United States. National Ocean Service. (2011). *The Gulf of Mexico at a Glance: A Second Glance.* Retrieved from National Oceanic and Atmospheric Administration: Institutional Repository: https://repository.library.noaa.gov/view/noaa/2674

United States. National Ocean Service. (2017). *2017 Ecosystem Status Report Update for the Gulf of Mexico.* Retrieved from National Oceanic and Atmospheric Administration: OCM Digitalcoast: https://www.aoml.noaa.gov/ocd/ocdweb/ESR_GOMIEA/

Valente, T. W. (2010). *Social Networks and Health: Models, Methods, and Applications.* New York: Oxford University Press.

Vallejo, J. (2021). *Evaluation of science collaboratives with the NOAA-WCR Central Valley California Office.* NOAA Center for Coastal and Marine Ecosystems (CCME).

Van den Besselaar, P., & Sandstrom, U. (2019). Measuring researcher independence using bibliometric data: A proposal for a new performance indicator. *PLoS One*.

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*, 523-538.

Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometrics networks. In Y. Ding, R. Rousseau, & D. Wolfram, *Measuring scholarly impact: Methods and practice* (pp. 285-320). Springer.

Van Eck, N. J., & Waltman, L. (2022, January 24). *VOSviewer - Getting Started.* Retrieved from VOSviewer: https://www.vosviewer.com/getting-started

Vera-Baceta, M., Thelwall, M., & Kousha, K. (2019). Web of Science and Scopus language coverage. *Scientometrics, 121*, 1803-1813.

Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications.* London: Cambridge Universityh Press.

Wondolleck, J. M., & Yaffee, S. (2017). Mobilizing a Multistate Partnership in the Gulf of Mexico. In S. L. Yaffee, & S. McKearnan, *Marine Ecosystem-based Managment in Practice: Different Pathways, Common Lessons* (pp. 43-73). Island Press.

Wondolleck, J. M., & Yaffee, S. L. (2017). *Marine ecosystem-based management in practice - Different pathways, common lessons.* Island Press.

Zaldivar-Jimenez, A., Ladron-de-Guevara-Porras, P., Perez-Caballos, R., Diaz-Mondragon, S., & Rosado-Solorzano, R. (2017). US-Mexico joint Gulf of Mexico large marine ecosystem based assessment and management: Experience in community involvement and mangrove wetland restoration in Términos lagoon, Mexico. *Environmental Develpment, 22*, 206-213.

APPENDIX

APPENDIX

STATA .DO FILES

NOTE #1: These .do files are included to make it possible to confirm this study's findings and investigate alternative specifications.

NOTE #2: Make sure that all files are located under one folder for analysis

STATA .do file: Data Management

```
cd "D:\GOM_Data"
clear
save population.dta, replace emptyok
clear
save publishing.dta, replace emptyok
clear
save rawpublishing.dta, replace emptyok
clear
save pairs.dta, replace emptyok
clear
*** Raw data ***
import excel using "GoM Final data - 6.3.2022 JV.xlsx", first clear
*** Uniform labels ***
fre Inst
replace Inst = "Inter-gov" if Inst=="inter-gov"
replace Inst = "Private" if Inst=="private"
fre Inst
sort Inst
fre Juris
replace Juris = "Inter-gov" if Juris=="inter-gov"
replace Juris = "International" if Juris=="INT"
replace Juris = "Cuba" if Juris=="CUB"
replace Juris = "Cuba Federal" if Juris=="CUB Federal"
fre Juris
*** Removal of duplicate organizations within each article ***
egen articleid = group(ArticleID)
sum articleid
```

```
save population0.dta, replace
forvalues d = 1/4603{
        use population0.dta, clear
        keep if articleid==`d'
        gen n = 1
        collapse (sum) n, by(Title Year Sourcetitle Aff Inst Country Cty_Inst Loc Juris)
        gen articleid='d'
        drop n
        append using population.dta
        save population.dta, replace
}
*** Removal of articles with new single entities ***
use population.dta, clear
bysort articleid: gen articleid_freq = _N
drop if articleid_freq < 2
drop if articleid_freq > 100
sort Year articleid
save population.dta, replace
***Generate ID numbers ***
        *Create SourceID for each article source remaining*
use population.dta, clear
egen SourceID = group(Sourcetitle)
sum SourceID
order SourceID, before(Sourcetitle)
save population.dta, replace
        *Create ArticleID for each article remaining*
use population.dta, clear
sort Year articleid
egen ArticleID = group(Title)
sum ArticleID
drop articleid
order ArticleID, before(Title)
save population.dta, replace
        *Create ID for each organization*
use population.dta, clear
egen AffID = group(Aff)
sum AffID
order AffID, before(Aff)
save population.dta, replace
        *Create ID number for Institution*
use population.dta, clear
egen InstID = group(Inst)
sum InstID
order InstID, before(Inst)
```

```
save population.dta, replace
        *Create ID number for Country Institution*
use population.dta, clear
egen Cty_InstID = group(Cty_Inst)
sum Cty_InstID
order Cty_InstID, before (Cty_Inst)
save population.dta, replace
        *Create ID for each country*
use population.dta, clear
egen CtyID = group(Country)
sum CtyID
order CtyID, before(Country)
save population.dta, replace
                *Create ID number for Location*
use population.dta, clear
egen LocID = group(Loc)
sum LocID
order LocID, before(Loc)
save population.dta, replace
        *Create ID number for Jurisdiction*
use population.dta, clear
egen JurisID = group(Juris)
sum JurisID
order JurisID, before(Juris)
save population.dta, replaceexport excel using
"D:\GOM_Data\GOM_STATA_Final.xlsx", firstrow(variables) replace
```

STATA .do file: Publication Analysis

```
cd "D:\GOM_Data"
*** number of articles published per source from R.A.W. data ***
clear
save rawpublishing.dta, replace emptyok
import excel using "GoM SCOPUS Master File RAW.xlsx", first clear
sort Year
egen SourceID = group(Sourcetitle)
drop if SourceID == .
sort SourceID
sum SourceID
```

```
sort Year SourceID
egen ArticleID = group(Title Sourcetitle)
sum ArticleID
save rawpublishing.dta, replace
keep Title Year Sourcetitle SourceID ArticleID
export excel using "D:\ GOM_Data\GOM_STATA_RAW_Articles.xlsx",
firstrow(variables) replace
        *Frequency of articles per source*
use rawpublishing.dta, clear
tabulate Sourcetitle, sort
        table ( Sourcetitle ) () (), statistic(frequency) statistic(percent)
        collect export "gomrawsourcefre", as(xlsx) sheet(Sheet1) cell(A1) replace
        *Frequency of articles per year*
use rawpublishing.dta, clear
tabulate year, sort
        table ( Year ) () (), statistic(frequency) statistic(percent)
        collect export "gomrawyearfre", as(xlsx) sheet(Sheet1) cell(A1) replace
*** Number of final publications per analysis focus ***
use population.dta, clear
save publishing.dta, replace
        * affiliation *
use publishing.dta, clear
bysort AffID: generate Aff_Publications=_N
order Aff_Publications, after(Aff)
tabulate Aff, sort
        table ( Aff ) () (), statistic(frequency) statistic(percent)
        collect export "affperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        * institutions *
bysort InstID: generate Inst_Publications=_N
order Inst_Publications, after(Inst)
tabulate Inst, sort
        table ( Inst ) () (), statistic(frequency) statistic(percent)
        collect export "instperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        * country *
bysort CtyID: generate Cty_Publications=_N
order Cty_Publications, after(Country)
tabulate country, sort
        table ( Country ) () (), statistic(frequency) statistic(percent)
        collect export "ctyperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        *location *
bysort LocID: generate Loc_Publications=_N
order Loc_Publications, after(Loc)
tabulate Loc, sort
        table ( Loc ) () (), statistic(frequency) statistic(percent)
```

```
	collect export "locperart", as(xlsx) sheet(Sheet1) cell(A1) replace
	* jurisdiction *
bysort JurisID: generate Juris_Publications=_N
order Juris_Publications, after(Juris)
tabulate Juris, sort
	table ( Juris ) () (), statistic(frequency) statistic(percent)
	collect export "jurisperart", as(xlsx) sheet(Sheet1) cell(A1) replace
save publishing.dta, replace
export excel using "D:\GOM_Data\GOM_STATA_Publishing.xlsx",
firstrow(variables) replace
*** number of articles published per source and by year ***
	*Removal of secondary+ authors in addition to other variables for ease of
analysis*
use publishing.dta, clear
sort Year ArticleID
bysort Year ArticleID : generate AuthorbyArticleID = _n
drop if AuthorbyArticleID > 1
sort SourceID
keep Title Year Sourcetitle SourceID ArticleID
	*Frequency of articles per source*
tabulate Sourcetitle, sort
	table ( Sourcetitle ) () (), statistic(frequency) statistic(percent)
	collect export "sourcefre", as(xlsx) sheet(Sheet1) cell(A1) replace
	*Frequency of articles per year*
tabulate year, sort
	table ( Year ) () (), statistic(frequency) statistic(percent)
save articles.dta, replace
export excel using "D:\ GOM_STATA_Articles.xlsx", firstrow(variables) replace
***Gulf of Mexico publication focus***
use population.dta, clear
save gompublishing.dta, replace
use gompublishing.dta, clear
drop if JurisID==17
drop if JurisID==7
drop if JurisID==10
save gompublishing.dta, replace
	* affiliation *
use gompublishing.dta, clear
bysort AffID: generate Aff_Publications=_N,
order Aff_Publications, after(Aff)
tabulate Aff, sort
	table ( Aff ) () (), statistic(frequency) statistic(percent)
	collect export "gomaffperart", as(xlsx) sheet(Sheet1) cell(A1) replace
	* institutions *
```

```
bysort InstID: generate Inst_Publications=_N
order Inst_Publications, after(Inst)
tabulate Inst, sort
        table ( Inst ) () (), statistic(frequency) statistic(percent)
        collect export "gominstperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        * country jurisdiction *
bysort Cty_InstID: generate Cty_InstPublications=_N
order Cty_InstPublications, after(Cty_Inst)
tabulate Cty_Inst, sort
        table ( Cty_Inst ) () (), statistic(frequency) statistic(percent)
        collect export "ctyinstperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        *location *
bysort LocID: generate Loc_Publications=_N
order Loc_Publications, after(Loc)
tabulate Loc, sort
        table ( Loc ) () (), statistic(frequency) statistic(percent)
        collect export "gomlocperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        * jurisdiction *
bysort JurisID: generate Juris_Publications=_N
order Juris_Publications, after(Juris)
tabulate Juris, sort
        table ( Juris ) () (), statistic(frequency) statistic(percent)
        collect export "gomjurisperart", as(xlsx) sheet(Sheet1) cell(A1) replace
        save gompublishing.dta, replace
export excel using "D: \GOM_Data\GOM_STATA_GOMPublishing.xlsx",
firstrow(variables) replace
        *Removal of secondary+ authors in addition to other variables for ease of
analysis*
use gompublishing.dta, clear
sort Year ArticleID
bysort Year ArticleID : generate AuthorbyArticleID = _n
drop if AuthorbyArticleID > 1
sort SourceID
keep Title Year Sourcetitle SourceID ArticleID
        *Frequency of articles per source*
tabulate Sourcetitle, sort
        table ( Sourcetitle ) () (), statistic(frequency) statistic(percent)
        collect export "gomsourcefre", as(xlsx) sheet(Sheet1) cell(A1) replace
        *Frequency of articles per year*
tabulate year, sort
        table ( Year ) () (), statistic(frequency) statistic(percent)
        collect export "gomyearfre", as(xlsx) sheet(Sheet1) cell(A1) replace
save gomarticles.dta, replace
export excel using "D:\GOM_STATA_GOMArticles.xlsx", firstrow(variables) replace
```

STATA .do file: Network Analysis

```
*** Organizations per variable***
use population.dta, clear
save organization.dta, replace
use organization.dta, clear
sorby Aff: aff_freq=_n
drop if aff_freq > 1
keep AffID Aff Country Loc Inst
save organization.dta, replace
export excel using "D:\GOM_Data\GOM_STATA_Organizations.xlsx",
firstrow(variables) replace
table ( Country ) () (), statistic(frequency) statistic(percent)
        collect export "orgpercty", as(xlsx) sheet(Sheet1) cell(A1) replace
table ( Loc) () (), statistic(frequency) statistic(percent)
        collect export "orgperloc", as(xlsx) sheet(Sheet1) cell(A1) replace
table ( Inst) () (), statistic(frequency) statistic(percent)
        collect export "orgperinst", as(xlsx) sheet(Sheet1) cell(A1) replace
*** Pairs within articles (links between two observations)***
use population.dta, clear
keep ArticleID AffID Aff InstID Inst CtyID Country lat lon LocID Loc JurisID Juris
save pairs.dta, replace
        *Find and remove duplicates*
use pairs.dta, clear
by ArticleID Aff, sort: gen i=_n
keep if i==1
drop i /*12513 observations*/
tempfile temp1
save `temp1', replace
        *create all possible pairs using joinby*
rename Aff b
rename AffID bx
rename InstID cx
rename CtyID dx
rename LocID ex
rename JurisID fx
rename Cty_InstID gx
joinby ArticleID using `temp1'
drop if Aff==b
egen d1=concat(AffID bx)
egen d2=concat(bx AffID)
```

```
replace d1=d2 if AffID>bx
          *Find and remove duplicate permutations within ArticleID*
by ArticleID d1, sort: gen i=_n
keep if i==1
drop d1
drop d2
drop i
drop if ArticleID==.
           * Frequencies/strength of each pair *
egen afffreq = count(1), by (bx AffID)
egen instfreq = count(1), by (cx InstID)
egen ctyfreq = count(1), by (dx CtyID)
egen locfreq = count(1), by (ex LocID)
egen jurisfreq = count(1), by (fx JurisID)
egen ctyinstfreq= count(1), by (gx Cty_InstID)
save pairs.dta, replace
export excel using "D:\GOM_Data\GOM_STATA_Pairs.xlsx", firstrow(variables)
replace

***Gulf of Mexico Geographical Focus***
use population.dta, clear
drop if JurisID==17
drop if JurisID==7
drop if JurisID==10
save gompopulation.dta, replace
use gompopulation, clear
keep ArticleID AffID Aff InstID Inst Cty_InstID Cty_Inst LocID Loc JurisID Juris
save gompairs.dta, replace
*** Affiliation/institution/jurisdiction pairs within articles (links between two
observations)***
          *find and remove duplicates*
use gompairs.dta, clear
by ArticleID Aff, sort: gen i=_n
keep if i==1
drop i
tempfile temp2
save `temp2', replace
          *create all possible pairs using joinby*
rename Aff b
rename AffID bx
rename InstID cx
rename LocID ex
rename JurisID fx
rename Cty_InstID gx
```

```
joinby ArticleID using `temp2'
drop if Aff==b
egen d1=concat(AffID bx)
egen d2=concat(bx AffID)
replace d1=d2 if AffID>bx
        *find and remove duplicate permutations within ArticleID*
by ArticleID d1, sort: gen i=_n
keep if i==1
drop d1
drop d2
drop i
        drop if ArticleID==.
* Frequencies/strength of each pair *
egen afffreq = count(1), by (bx AffID)
egen instfreq = count(1), by (cx InstID)
egen locfreq = count(1), by (ex LocID)
egen jurisfreq = count(1), by (fx JurisID)
egen ctyinstfreq= count(1), by (gx Cty_InstID)
save gompairs.dta, replace
export excel using "D:\GOM_Data\GOM_STATA_GOMPairs.xlsx",
firstrow(variables) replace
```

STATA .do file: Visualization Data Management

```
*NOTE1: Analysis Network files must be done prior to this one.*
*NOTE2: Repeat calculations using gompairs.dta file and saving new files under
different name*
 *** VOSViewer Nodes/Map files ***
use population.dta, clear
bysort Aff: generate idfreq = _n
drop if idfreq>1
drop idfreq
rename AffID Id
rename Aff Label
sort Id
drop if Id==.
keep Id Label
save affnode.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_affnode.csv", replace
```

```
	* institution *
use population.dta, clear
bysort Inst: generate idfreq = _n
drop if idfreq>1
drop idfreq
rename InstID Id
rename Inst Label
keep Id Label
sort Id
drop if Id==.
save instnode.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_instnode.csv", replace
	* countries *
use population.dta, clear
bysort Country: generate idfreq = _n
drop if idfreq>1
drop idfreq
rename CtyID Id
rename Country Label
rename lat y
rename lon x
keep Id Label x y
sort Id
drop if Id==.
order y, after(x)
save ctynode.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_ctynode.csv", replace
	* country institution*
use population.dta, clear
bysort Cty_Inst: generate idfreq = _n
drop if idfreq>1
drop idfreq
rename Cty_InstID Id
rename Cty_Inst Label
keep Id Label /
sort Id
drop if Id==.
save ctyinstnode.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_ctyinstnode.csv", replace
	* locations *
use population.dta, clear
bysort Loc: generate idfreq = _n
drop if idfreq>1
drop idfreq
```

```
rename LocID Id
rename Loc Label
keep Id Label
sort Id
drop if Id==.
save locnode.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_locnode.csv", replace
          * jurisdictions *
use population.dta, clear
bysort Juris: generate idfreq = _n
drop if idfreq>1
drop idfreq
rename JurisID Id
rename Juris Label
keep Id Label/
sort Id
drop if Id==.
save jurisnode.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_jurisnode.csv", replace
*** VOSViewer Network/Edge files ***
          *  affiliation *
use pairs.dta, clear
bysort bx AffID: generate idfreq = _n
drop if idfreq > 1
drop idfreq
rename AffID Source
rename bx Target
rename afffreq strength
sort ArticleID
keep Source Target Strength
save affedge.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_affedge.csv", replace
          * institution *
use pairs.dta, clear
bysort cx InstID: generate idfreq = _n
drop if idfreq > 1
drop idfreq
rename cx Source
rename InstID Target
rename instfreq strength
sort ArticleID
keep Source Target Strength
save instedge.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_instedge.csv", replace
```

```
	* countries *
use pairs.dta, clear
bysort dx CtyID: generate idfreq = _n
drop if idfreq > 1
drop idfreq
rename dx Source
rename CtyID Target
rename ctyfreq strength
sort ArticleID
keep Source Target Strength
save ctyedge.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_ctyedge.csv", replace
	* country institution *
use pairs.dta, clear
bysort gx Cty_InstID: generate idfreq = _n
drop if idfreq > 1
drop idfreq
rename gx Source
rename Cty_InstID Target
rename ctyinstfreq strength
sort ArticleID
keep Source Target Strength
save ctyinstedge.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_ctyinstedge.csv", replace
	* locations *
use pairs.dta, clear
bysort ex LocID: generate idfreq = _n
drop if idfreq > 1
drop idfreq
rename ex Source
rename LocID Target
rename locfreq strength
sort ArticleID
keep Source Target Strength
save locedge.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_locedge.csv", replace
	* jurisdiction *
use pairs.dta, clear
bysort fx JurisID: generate idfreq = _n
drop if idfreq > 1
drop idfreq
rename fx Source
rename JurisID Target
rename jurisfreq strength
```

```
sort ArticleID
keep Source Target Strength
save jurisedge.dta, replace
export delimited using "D:\GOM_Data\GOM_STATA_jurisedge.csv", replace
```

BIOGRAPHICAL SKETCH

Juliet S. Vallejo graduated from Texas A&M University in 2013, receiving a Bachelor of Science in Wildlife and Fisheries Science with a focus in wildlife ecology and conservation. She later obtained her Master of Science in Ocean, Coastal, and Earth Science at The University of Texas Rio Grande Valley and graduated in August 2022. While in her graduate studies, she also completed a NOAA Experimental Research & Training Opportunity internship for the NOAA Fisheries West Coast Region California Central Valley Office. Before her master's studies, Juliet interned with U.S. Fish and Wildlife Service and later worked with Texas Parks and Wildlife Department and the Texas A&M Agrilife Research Center. She can be reached at jvallejo1019@gmail.com.