Visualizing Network Structures in the Food, Energy, and Water Nexus

by

Brandon Mathis

A Thesis Presented in Partial Fulfillment
of the Requirement for the Degree
Master of Science

Approved February 2019 by the
Graduate Supervisory Committee:

Ross Maciejewski, Chair
Giuseppe Mascaro
Hessam Sarjoughian

ARIZONA STATE UNIVERSITY

May 2019

ABSTRACT

In recent years, the food, energy, and water (FEW) nexus has become a topic of considerable importance and has spurred research in many scientific and technical fields. This increased interest stems from the high level, and broad area, of impact that could occur in the long term if the interactions between these complex FEW sectors are incorrectly or only partially defined. For this reason, a significant amount of interdisciplinary collaboration is needed to accurately define these interactions and produce viable solutions to help sustain and secure resources within these sectors. Providing tools that effectively promote interdisciplinary collaboration would allow for the development of a better understanding of FEW nexus interactions, support FEW policy-making under uncertainty, facilitate identification of critical design requirements for FEW visualizations, and encourage proactive FEW visualization design.

The goal of this research will be the completion of 3 primary objectives: (i) specify visualization design requirements relating to the FEW nexus; (ii) develop visualization approaches for the FEW nexus; and (iii) provide a comparison of current FEW visualization approaches against the proposed visualization approach. These objectives will be accomplished by reviewing graph-based visualization, network evolution, and visual analysis of volume data tasks, discussion with domain experts, examination of currently used visualization methods in FEW research, and conduction of a user study. This will provide a more thorough and representative depiction of the FEW nexus, as well as a basis for further research in the area of FEW visualization. This research will enhance collaboration between policymakers and domain experts in an attempt to encourage in-depth nexus research that will help support informed policy-making and promote future resource security.

To my family and friends for all of their encouragement and support

ACKNOWLEDGMENTS

I would first like to express my gratitude to my advisor, Dr. Ross Maciejewski, for giving me the opportunity to work in his lab and for his genuine concern about the success of his students. Without his knowledge and guidance this thesis would not have been possible. I would also like to give thanks to Dr. Giuseppe Mascaro and Dr. Hessam Sarjoughian for being part of my graduate committee and for providing feedback and recommendations throughout my research.

Furthermore, I would like to thank my colleagues in the Visual Analytics and Data Exploration Research (VADER) Lab for their criticism and suggestions regarding this work. I would also like to thank Professor Ted LaRue from Fairmont State University for his continual encouragement to better myself at everything I do, and for pushing me to attend graduate school. Finally, I would like to thank my family and friends for their support throughout this process. Most importantly, I would like to thank my mom for her unwavering support through everything that I pursue and for the countless hours she has spent helping with revisions.

This work was supported in part by the National Science Foundation under Grant No. 1639227 and the U.S. Department of Homeland Security under Award Number, 2017-ST-061-QA0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

TABLE OF CONTENTS

APPENDIX

LIST OF FIGURES

Chapter 1

INTRODUCTION

Previous literature that examines the food, energy, and water (FEW) nexus often focuses on only one or two parts of the nexus rather than all of the parts as a whole. Historically, this literature also does not investigate how climate can affect the nexus, although it is generally regarded as an element that also needs to be considered (Berardy and Chester, 2017). Viewing the nexus as one unified system is necessary, as changes to one of the parts often impacts the others as well. Without representing all of these impacts within the system, much of the information gained can be skewed and may show the system in a state that is not an accurate representation of what could really occur.

Attempting to visualize the interactions between the sectors in the FEW nexus is not a new task, but it is one that requires significant research. Currently, little research has been done in the area of FEW nexus visualizations. This poses a problem for the development of new, or even hybrid, visualization methods. In terms of current visualizations, there are a number of challenges and limitations that appear when attempting to use them for FEW Data Visualization, such as order of magnitude differences in usage between the FEW sectors and granularity of the data available. This is due to the complexity of the interactions that occur within the FEW nexus, as well as an inability to accurately define all of these complex interactions in full.

In most current FEW nexus research, interactions between the sectors in the nexus are generally defined or presented at a high level. Often, this means that a sector is abstracted and represented as a single component rather than multiple components.

This removes a large amount of detail that is often necessary for effective policy decisions and for the examination of potential future scenarios that could occur.

The visualization approach that is often used for FEW visualization, or at least Energy-Water visualization, is the Sankey diagram. The general layout of the Sankey diagram was first introduced in Minard (1862), Charles Minard's Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812-1813. The problem is not whether Sankey diagrams are a good visualization approach but whether Sankey diagrams are the best way to visualize the highly complex FEW nexus. The question at hand is whether Sankey diagrams are currently only used because that is what has traditionally been done and implementing a new design would have a high learning cost. Soundararajan *et al.* (2014) state that Lawrence Livermore National Laboratory (LLNL) has been creating energy flow Sankey diagrams since 1976. This shows a large amount of work has gone into developing these diagrams and using a new visualization approach may be a difficult adjustment, although new visualizations could potentially be beneficial in completing a number of tasks.

In this work, three visualization design alternatives were developed to encourage collaboration between policymakers and domain experts in an attempt to promote in-depth nexus research that will help to better understand these interactions and secure future resource security within the nexus. More specifically, this work proposes:

1. **Visualization design requirements** relating to the FEW nexus developed through:
    (a) Review of graph-based visualization task taxonomies
    (b) Review of network evolution task taxonomies
    (c) Review of visual analysis of volume data task taxonomies

    (d) Analysis of current energy-water diagrams produced by Lawerence Livermore National Lab
    (e) Discussions with domain experts

2. **Visualization design alternatives** that attempt to improve readability and minimize the negative impact of edge crossings through the incorporation of characteristics from:
    (a) Sankey diagrams
    (b) Treemaps
    (c) Graphs/Networks

This work shows that, although not generally explored as a solution to a visualization problem, the development of new visualization designs can be positively received by users and can provide detailed information needed for accurate analysis that is not available using traditional visualization approaches.

The remainder of this thesis is organized as follows. Chapter 2 introduces the FEW nexus and approaches for the narration of data. Chapter 3 presents other works related to uncertain future analysis, task analysis, visualization design aesthetics, and visualization approaches in critical infrastructure. Chapter 4 gives design requirements for a FEW visualization and the development leading to the final visualization designs. Chapter 5 presents the user study design and results obtained. Chapter 6 presents limitations in the user study and limitations of flow diagrams for use in the FEW nexus. Chapter 7 presents the formal conclusions and provides recommendations for future work.

Chapter 2

VISUALIZING NEXUS THINKING

2.1 FEW Management

When looking at how to manage FEW, or other, resources, the term policy lever is often used. A policy lever is a potential policy decision that could be made by a policymaker. A number of potential policy levers that can be considered when looking at policies relating to water were identified by Olmstead (2014), including price, diversion of resources to different sectors, per capita use, reclaimed water usage, river flow change, and use restrictions. Many of these potential levers can easily be adapted to the food and energy sectors as well. Giving a universal list of policy levers for all applications, and for all systems, is not possible, as deciding on which of these levers to pursue is highly dependent on the area and the control that the policy maker has within that area. When implementing policies within urban areas, there are generally three main categories of conservation that are explored. These categories are adoption of conservative technologies, use restrictions, and social comparison and information policies (Olmstead, 2014).

In recent years, a strong effort, mainly through new policies, has been made to improve energy efficiency and promote the use of renewable energy sources across the United States. According to Bartos and Chester (2014), this has led to the need for further exploration of the water and energy nexus, as gaining a better understanding of the interdependencies between these sectors could significantly increase energy efficiency, reduce water usage, and help shape future policy decisions. As water is imperative for the production of food and other crops, water sustainability is an important issue, mainly within the American Southwest and other similar regions. This is mostly due to the population rapidly growing, agricultural demand increasing due to rising temperatures, and the limited supply of groundwater. Emergency restrictions on outdoor water use have already been placed on many large cities in this area, including Los Angeles, Albuquerque, and Tucson (MacDonald, 2010). This reinforces the fact that water policy decisions are critical within these areas, and all water policy decisions under different scenarios need to be examined carefully.

When examining critical infrastructure systems, the issue of scale can often arise. This is due to the management of water resources being examined at a local or regional scale, as this is the main area of impact. Adversely, the impact of energy resources can be examined at a global scale, as the effect of things, such as emissions, can impact the world as a whole, rather than just a single region. Not only does the issue of scale occur across different sectors, but examining a single sector at different scales can result in different characteristics of an area (Curmi *et al.*, 2013). This discrepancy in scale can lead to issues in presenting meaningful results to end users depending on their role in resource management.

Due to the varying roles of end users that might interact with the data being presented, the issue of framing also becomes prominent. Hoolohan *et al.* (2018) state that nexus issues can broadly be framed as security, dependency, scarcity, or risk. Without a clear understanding of the intended framing of the nexus, a visualization may not present an appropriate solution to the end user. Stakeholder engagement is a large part of nexus research, as the problem definition can be better established and more appropriate solutions can be offered (Hoolohan *et al.*, 2018). However, the vast variation in end users involved in FEW research and policy also creates a scientific language barrier problem. In a work by Mattor *et al.* (2014), they examine transdisciplinary language barriers in environmental governance. Their work required the development of a shared understanding of concepts and a shared language, which took a significant amount of time due to the diverse backgrounds of the users involved. This is an issue within the FEW nexus as there are many different disciplines that are involved, as well as variations within the disciplines themselves.

2.2 Facilitating Interaction with Data

Narration has long been the primary means of presenting and sharing information between individuals. Although narration has not traditionally been used in data visualization, it can be beneficial for a number of reasons. Segel and Heer (2010) state that when looking at narration, narrative visualizations can generally be classified somewhere within an author-driven to reader-driven approach spectrum. Figueiras (2014) determined that approaching visualization from a narrative perspective provides the opportunity for information to be presented in a structured context, rather than left for the end user to discern. Using narration can also help to keep the user engaged with the information that is being presented, as the user is able to explore the different areas that interest them, rather than just being presented a large amount of data that may seem incoherent.

When designing any visualization or decision support system, it is imperative to understand what information the user needs available to them and how they plan to interact with that information. This is especially true when it comes to designing applications to be used by policymakers, as they need to plan and make decisions that will have an impact on a large number of people. For this kind of application, a user-centered design needs to be used. In a work by Hadhrawi *et al.* (2013), the authors develop a web-based collaborative planning interface platform, CoPI, and perform a case study to show its efficacy in policy planning of large-scale infrastructure. From this study, it is easy to see that developing an application with a user-centered design approach requires a significant amount of stakeholder engagement but also gives a better insight into the way that the stakeholders expect the application to function.

Humans possess a number of important skills, many of which are desirable skills to reproduce in a computer for a number of tasks. Green *et al.* (2008) identify two of these, adaptation and accommodation of information. These two skills are the basis of human learning and perception of new and complex information. Through adaptation and accommodation, low-level connections between new information and similar information that has been previously learned are formed. This type of categorization of information also allows humans to be able to discern information that is of little or no importance to their current task relatively quickly (Green *et al.*, 2008).

Green *et al.* (2008) further identify a key weakness of a human when compared to a computer is the lack of a working memory, or a memory that has pseudo-infinite capacity. This allows computers to maintain all pertinent information, whereas a human only has the ability to remember a finite amount of information. Humans also suffer from the inability to retain focus on previous information when presented with new information. Thus, when dealing with complex systems, a Human Cognition Modeling approach is often employed. This approach allows the user to select information of interest and the computer to facilitate the presentation of that and related information.

A similar approach is used in a study by Quintero *et al.* (2005), where a Case-Based Reasoning interactive decision support system (IDSS) was developed. They determined that in this kind of IDSS, the exchange between the user and the IDSS can be simply stated as the posing of questions from the IDSS to a user, the responses from the user to the IDSS, and a set of potential actions returned from the IDSS to the user. These questions have the general form: what is the problem, where is the problem, and how did the problem occur. This helps to narrow down the search space for the IDSS, as it now knows generally what part of the system the problem was occurring and what event or attribute may have triggered it. The IDSS then returns a list of selected potential actions and the user decides whether they want to choose one of the recommended actions or one of their own. This approach gives the user full control of the decision-making process while also suggesting alternatives or actions that may not have been considered.

To help highlight the potential impact of recommendations given by decision support systems, a number of different techniques can be utilized to aid in the narration of pertinent data to an end user. Wang *et al.* (2010) determined that when displaying data that uses similar axis values or related data sets, small multiple views can be used to allow the user to easily examine the differences between similar items. Having this information displayed in a single view also reinforces the coherency of the information that is being presented. This also presents the opportunity to highlight data that may be of interest to the user by representing positive or negative scenarios with different colors. Although accuracy is a large factor when dealing with any kind of intelligent system, this could aid in drawing attention to data that has the potential to be of importance to a user (Wang *et al.*, 2010).

Another potential technique, in web-based data visualization, is to employ a single page web application. Tesarik *et al.* (2008) allude that designing these kind of applications, however, requires careful planning by the developer so that all of the important information that is to be displayed or manipulated can be accessed in an intuitive way. Since a single web page provides a limited amount of space, other methods have to be used to present more detailed information about a topic. This can be done by using tooltips or modal popups, much like what would be seen in a desktop application. Approaching the design in this way can help to improve the user experience, as the way that the information is presented to the user promotes a sense of connection between all of the information of the web page (Tesarik *et al.*, 2008).

Through the inclusion of similar narration techniques, visual analytics systems can be used to present a more comprehensive view of the underlying data. This makes considering the elements that could be used for highlighting within a visualization design an important part of the visualization design process. If elements are too low-level, it may cause the visualization to not be easy to understand and if elements are too high-level, it might not provide enough information to get an accurate depiction of the data. Taking visual analytics into account can also help to shape the overall visualization design. Knowing that an end user is expected to interact with data in a certain order can structure the ordering of the visualization design itself. This can simplify the interactions of an IDSS by making transitions from one piece of data to the next piece of data shorter or more intuitive, like moving from the first stage to the second stage of a flow diagram.

Chapter 3

RELATED WORK

3.1 Uncertain Future Analysis

In a work by Sullivan *et al.* (2017), water sensitivity in the Colorado River Basin was examined by looking at historical water governance transitions. In this comparative historical water governance approach, key water governance transitionary events occurring from the early 1800s through 2011 were examined and discussed to identify historical patterns and governance that may impact or limit future governance transitions. It was concluded that water governance within the western United States is distinct from water governance in the rest of the country, and that historically, meaningful water governance changes within this area can take decades of work to implement. By drawing from historical data, a better understanding of the complexity of governance within a region can be formed and can allude to the likelihood of the acceptance of future governance transitions.

In a study by Berardy and Chester (2017), climate change vulnerability in the FEW nexus was examined and climate was determined to be a significant factor in the security of the other systems, especially in terms of Arizona's agriculture. Alfalfa and cotton are two of Arizona's most plentiful crops, together making up approximately 51% of the total crop acreage. The problem is that alfalfa and cotton also require the largest amounts of water. Not only that, but an experiment looking at the effects of summer irrigation termination on alfalfa in Tucson, Arizona showed that plant mortality increased by between 25% to 35% when subject to discontinued irrigation for more than two months (Berardy and Chester, 2017). Because of Arizona's climate and increasing population, the energy and water systems are already strained during the summer months. Adding increased irrigation demand for agriculture could result in interruptions that would inevitably result in even more irrigation demand for agriculture. This would not only have a negative impact on Arizona's crops but could also cause issues in other major urban centers in the Southwest, such as El Paso and Los Angeles (Berardy and Chester, 2017).

Similarly, a case study by Gober and Kirkwood (2010) in Phoenix, Arizona was proposed in which a simulation, called WaterSim, of long-term water shortage was examined. Through the implementation of policy levers within WaterSim, a number of varying analyses could be performed not only by experts in the field, but also by a number of other end users. This policy lever approach also yielded results that could realistically be achieved by policy makers' decisions rather than focusing on producing ideal results that may not be attainable. Not only did this study provide a valuable simulation model for examining the potential impact of different policies under varying future conditions, it also reinforced the need for interactive systems to aid in policy decisions.

In environmental governance, due to the highly interdisciplinary nature of this area, collaborative tools can often be a valuable resource to help collaborators share their expertise and provide domain specific insights to other members of their group that may not otherwise have been discerned from the underlying data. The problem with a number of these tools is that they require installation of specific toolkits or only allow for basic multi-user interactions, which can limit the level of collaboration and accessibility (Lukasczyk *et al.*, 2015). In an attempt to mitigate these issues, a web-based collaborative Data Visualization and Analysis framework was developed by Lukasczyk *et al.* (2015). This framework allowed for collaboration among users within a Visualization Room and provided the ability to share snapshots of visualizations performed by an individual in the Visualization Room with the other members of the group. The feedback from the users of this system was overall positive and further validated that collaborative systems are a good tool when attempting to solve complex problems.

As critical infrastructure systems are highly complex systems, they are extremely difficult to identify interdependencies between. This difficulty can further be compounded by the fact that these systems are highly adaptive and are often changed due to either past experiences or plausible assumptions of future demands. In recent years, a greater appreciation of the importance of the interdependencies in these systems has emerged, and has thus propelled research in this area (Rinaldi *et al.*, 2001). As this type of research is highly interdisciplinary, experts from a wide variety of domains are often needed to be able to develop models of critical infrastructure systems correctly. When attempting to develop a comprehensive framework for modeling and simulation of these systems, often three things are needed: access to a large amount of data, legacy or specifically developed software, and some set of evaluation metrics (Rinaldi *et al.*, 2001).

3.2 Task Analysis

According to Keim *et al.* (2008), Visual Analytics is a highly interdisciplinary field that makes use of visualization, data mining, data management, data fusion, statistics, and cognition science, and which relies heavily on interactive visualizations to facilitate understanding of large, complex data sets. As such, Visual Analytics draws a number of presentation techniques from traditional narration, like those presented in Section 2.2. Since Visual Analytics has a strong foundation in decision support systems, identifying low-level task requirements that are often used in these systems provides requirements for designing new visualizations, as well as a basis for evaluation metrics for these visualizations.

Task taxonomies have become a popular means of identifying a set of characteristics that help to validate the effectiveness of a visualization approach in terms of common exploratory tasks. This is done through the abstraction of interactions between the user and a visualization. These interactions are generally grouped into either low level-tasks, like those proposed by Amar *et al.* (2005), or high-level tasks, like those proposed by Ahn *et al.* (2014).

In a work by Lee *et al.* (2006), they focused on graph-based visual analytics tasks and identified the four major task categories: topology-based, attribute-based, browsing, and overview. Several of these task categories, as well as task categories in other task taxonomies, follow the Visual Information-Seeking Mantra proposed by Shneiderman (1996). This mantra proposes the three key concepts: overview, zoom and filter, and details-on-demand. Although this is not a definitive guideline for data visualization, this mantra has become a point of consideration when developing and examining visualization designs.

In the works by Ahn *et al.* (2014) and Kerracher *et al.* (2015), temporal tasks were the main area of focus for the analyses. Ahn *et al.* (2014) produced a set of higher level task categories which were derived from a set of fifteen real systems and studies, which included temporal data processing tasks, individual temporal event tasks, aggregated temporal event tasks, rate of changes, and compound tasks. The tasks identified by Kerracher *et al.* (2015) were developed from the review of 95 papers, including the work by Ahn *et al.* (2014), and contained the three main categories: structural, attribute based in a structural context, and attribute based. Both of these works highlight that the categorization of tasks can be difficult and often tasks can fall into more than one category.

The tasks associated with volume data are often different than the tasks that are generally identified for graph-based visualization and network evolution. This is due to volume visualization focusing on the representation of an object rather than how elements are connected to each other. In a work by Laha *et al.* (2015), they focused on the classification of user tasks relating to visual analysis in volume data. In this work, the task categories identified were search, pattern recognition, spatial understanding, quantitative understanding, and shape descriptor.

In a study by Dimara *et al.* (2018), conceptual and methodological issues that arise in multidimensional visualizations for decision support systems were analyzed. This study focused on multi-attribute choice tasks which highlighted a number of issues that could be encountered, such as the lack of a “right” answer for many multi-attribute choice tasks. To identify these issues, they evaluated parallel coordinates, scatterplot matrix, and tabular visualizations using both subjective and objective metrics. Both subjective and objective metrics were considered because although good decisions require a good understanding of the data at hand, having a good understanding of the data does not necessitate good decisions. The objective performance metrics focused on accuracy and time-on-task, and the subjective performance metrics focused on technique preference and choice assessment. They state that, although their work only focused on elementary and generic visualization methods, it could still be used to help produce evaluation methods for more complex visualizations as well.

Gotz and Zhou (2008) identify common visualization actions by reviewing Visual Analytics systems in an attempt to define an action taxonomy. The defined taxonomy contained three top-level categories: exploration actions, insight actions, and meta actions. This taxonomy was then used to develop a collaborative web-based visual analytic system, HARVEST. HARVEST allowed collaborative visual exploration of data and provided the user a record of the previous actions that were performed. HARVEST was well received by both developers, in terms of ability to implement new visualization widgets, and analysts, who stated that the displayed actions matched well with what they considered to be the steps that they took. The taxonomy that was identified in this work can help guide the development of new action types by providing a set of common visualization actions that can be expanded upon for specific visualization areas.

Although identifying general task taxonomies is beneficial in providing a solid basis for the further development of specific task taxonomies, Miksch and Aigner (2014) take a slightly different approach. They introduce the idea of a design triangle, in which they consider the data, the user, and the tasks of the users. Contrary to the other works in this section, this approach yields tasks for a specific situation with specific end users in mind rather than a generalized set of tasks. In their study, they apply the design triangle to three examples and produce the tasks, exploring the effects of clinical actions on patient's conditions, exploring organizational changes over times, and time series model selection. Although generally very high-level tasks, the identification of these kind of tasks can help to frame the visualization in such a way that the necessary information can be presented to an end user in a manner that promotes exploration.

Task taxonomies give insight into the potential uses of a visualization design and can facilitate the development of elements that can help to support the design's desired use cases. As task taxonomies cannot be specified for all applications of all visualization designs, task taxonomies for designs that share similar characteristics to a proposed design can be used as a basis for what might also be considered useful in a proposed design. Although many of the previously mentioned task taxonomies appear disconnected, many of the characteristics of the visualizations designs associated with these taxonomies are also shared with Sankey diagrams. The use of these taxonomies also yields a form of evaluation to assess the overall usefulness of a visualization design for its target application, or specific use case.

3.3 Visualization Design Aesthetics

Visualization design aesthetics are important for consideration when attempting to design or optimize the layout of design elements. Numerous papers have offered strategies and suggestions for improving automatic visualization layouts. In the work by Tanahashi and Ma (2012), the authors attempt to optimize the layout of storyline visualizations. In this work, they seek to minimize edge wiggles, edge crossings, and whitespace gaps. Similarly, the design criteria proposed by Purchase *et al.* (2001) includes: minimization of crossings between connections, minimization of the global length of connections, minimization of the area of the smallest rectangle covering the diagram, and placement on the external boundary of symbols representing interfaces. Rather than focusing on improving these diagrams, Purchase *et al.* (2001) focuses on which aesthetic elements should be considered as a priority through preference experiments. In terms of the UML class diagrams examined in this study, preference for fewer edge crossings and fewer bends were the most preferred aesthetic choices, with preferences of 93% and 91% respectively. UML class diagram aesthetics have also been considered by Eichelberger (2003) and include a number of edge focused criteria. These works align with the general consensus that edge crossings and sharp bends can have a strong impact on a visualization design. Similar preference focused studies have been performed by Harrison *et al.* (2015) and Inbar *et al.* (2007) in the areas of infographic and information visualization design aesthetics respectively.

In the work by Batini *et al.* (1986), they evaluate the six aesthetic criteria: minimize bends, minimize edge crossings, orthogonality, width of layout, text direction, and font type. Although not as recent as many of the other works presented here, the algorithm proposed by Batini *et al.* (1986) still serves as a reasonable point of consideration. Their algorithm uses the general steps: planarization, orthogonalization, and compaction, which attempt to maintain their six aesthetic criteria, although some compromises have to be made during the planarization step. Cui *et al.* (2008) similarly attempt to reduce the impact of edge crossings on the overall design. They accomplish this by introducing geometry-based edge bundles that seek to have edges meet at singular control point, rather than allowing a large number of crossings throughout the entire diagram. Other edge bundling based solutions can be seen in the works of Selassie *et al.* (2011), Holten and Van Wijk (2009), Gansner *et al.* (2011), Zhou *et al.* (2013), and Luo *et al.* (2012).

Work by Stasko and Zhang (2000) examines improving radial, space-filling hierarchical visualizations. The primary goal of this work was to allow small items to be viewed in detail, that would otherwise not be available, through the addition of interaction. They pose the design guidelines: maintain a full circular space-filling methodology and global view of the entire hierarchy, allow a more detailed examination of small peripheral files and directories, but keep it in context of the entire information structure, ensure that the overview of the entire hierarchy remains relatively stable in layout, maintain a balance between the visibility of both the hierarchy overview and the detailed focus display, be more space-efficient, avoid the use of multiple windows or scrollbars, and allow the viewer to easily track display changes between a global view and a detailed view of the peripheral nodes. To meet these design guidelines they pose the 3 methods: angular detail, detail outside, and detail inside. These methods allowed for smaller elements to be presented at their correct sizes, relative to the rest of the diagram, but also allowed for interaction to be used to magnify the elements for better detail. This highlights the usefulness of interaction when performing data visualization and shows that difficulty may be faced when attempting to present this kind of information using only a static visualization.

Flow diagrams are a common visualization design in many application areas. In terms of critical infrastructure and manufacturing, Sankey diagrams are generally the flow diagram that is utilized. These diagrams, however, have a number of challenges that have been examined and improved on in a number of works. In a work by Alemasoom *et al.* (2016), they attempt to mitigate some of the more common issues faced in Sankey diagrams by adding artificial, intermediate, nodes. This was done to create more aesthetically pleasing Sankey diagrams based on their criteria to minimize edge crossings, create edge lengths that are as short as possible, and maintain straight edges. Similar aesthetic improvements are sought in graphs by Breitkreutz *et al.* (2003), where they applied user-implemented node relaxation within their network visualization system, Osprey, in an attempt to reduce clutter and improve interpretability of complex networks. Although Osprey supports very simple layouts, many of these layouts do help to significantly reduce clutter that could be seen in networks that have a large number of nodes.

Graph layouts can become significantly more difficult as the number of edges in a graph increases. In a work by Martin *et al.* (2011), they developed an open-source toolbox, OpenOrd, to handle large graph layouts that have a significant number of edges. They identify three problems that are associated with scaling a force-directed layout to a large graph: inability to correctly uncover the global structure, visually unappealing layouts, and a high running time. The developed method, based on the Frutcherman-Reingold algorithm for force-directed layouts, was applied to a Wikipedia dataset containing 659,388 nodes with 5,843,729 edges in which six levels of recursion were needed, as well as approximately four hours of processing time.

Although still quite difficult to comprehend visually, the network structure produced was much more comprehensible than the original network structure.

3.4 Visualizing Critical Infrastructure

Tolone (2009) determined that when performing an analysis of critical infrastructure systems, isolation of any one component of the system from the rest of the system can result in an incomplete or potentially invalid analysis of the overarching system, and the interactions contained within it. This is due to the dynamics of the system and how the behavior of the different components of the system are limited by the behavior of the other components of the system, not only by their own characteristics. The context of a system, which is difficult to convey when looking at only isolated components of the system, becomes more difficult when attempting to also represent the interactions between complex components in a meaningful way. Often, spatial analysis of a system is done by simply attempting to overlay important data from the system onto a map. Although a reasonably assumed approach, performing spatial analysis in this way can lead to a loss of how the interactions and dependencies between different components are occurring within the system. Instead, the focus of the analysis can be shifted to gaining a better conveyance of interactions and dependencies between the components, but this can result in a loss of spatial context which is also important in visualization (Tolone, 2009).

In this same study, Tolone (2009) identifies 4 main challenges that need to be addressed when performing Visual Analytics in critical infrastructure analysis. First, there needs to be a facilitation of reasoning in context. This means that interaction with the visualizations should occur at a level of meaning, rather than at a data level, as this imposes limits on which pieces of the data can be manipulated and ensures that changes to one component also occurs in all corresponding components.

The second challenge is that critical infrastructures are complex multi-dimensional systems that require not only technical analysis, but also behavioral analysis. This can result in collaboration between research areas which traditionally may not have occurred or may have occurred minimally. The third challenge is to enable cognitive integration by designing visualizations in such a way that there can be integrated analysis among connected activities within the system. Finally, there needs to be a high level of transparency of the analysis. Due to their complexity, validation of these kinds of systems may not be possible using traditional methods (Tolone, 2009).

Sankey diagrams are one of the primary visualization method used in the evaluation of a number of critical infrastructure systems, along with Parallel Coordinate Plots and Scatter Plots. In a work by Alemasoom *et al.* (2016), they examined the use of Sankey diagrams in energy systems and developed a visualization system, EnergyViz, to facilitate the exploration of Canada's energy system. As mentioned previously, a key addition from this work was the use of artificial, intermediate nodes in their Sankey diagrams. Another key takeaway from this work was the use of level-of-detail exploration. This allowed for more detail to be provided to experienced end users while also simplifying the overview of the Sankey diagram for individuals that may not have previous experience with them. The overall feedback of the level-of-detail exploration was positive and highlighted that this approach could be beneficial when dealing with diverse groups of individuals. The final key addition from this work was the use of 3D Sankey diagrams. In this approach, a 2D Sankey diagram is used to present the flow of resources through the system, and bar charts are placed perpendicular to the plane of the 2D Sankey diagram to represent the different emissions at each stage of the flow (Alemasoom *et al.*, 2016). Overall, EnergyViz was well received by both experts and non-experts, and presents a number of improvements to Sankey diagrams for use in energy system analysis.

In a case study by Yang and Wi (2018), an interactive web-based visualization was developed to examine the FEW nexus in the Great Ruaha River of Tanzania. The underlying visualizations used in this study were a dynamic scatter plot matrix and a dynamic parallel coordinates plot. A dynamic scatter plot matrix was useful in this case study, as it provided a means to better understand the interactions between different management concerns, without requiring the user to have a technical background. A dynamic parallel coordinates plot was useful, as it could be used to represent connections between the different management objectives within the system, and because parallel coordinate plots are already used in water resource system analysis. The use of these visualizations allowed for a quick analysis of different scenarios and objectives without the need to individually analyze them for the specific desired characteristics (Yang and Wi, 2018).

Visualization within critical infrastructure is a difficult task, but is a task that is necessary to support knowledge exchange between domain experts and policy makers. In order for visualization to effectively facilitate this knowledge exchange, a level of trust in the data presented has to be established and the data needs to be presented in a way that is common to the domain at hand. As flow diagrams are commonly seen in critical infrastructure already, some form of flow diagram is a viable option for representing the FEW nexus to domain experts in the areas of food, energy, and water systems. Flow diagrams should appeal to policy makers as well because of the sequential nature of the design as well as its explicit representation of inputs and outputs to the system.

Chapter 4

SANKEY DESIGN

4.1 Design Requirements

Little work has been done in the specification of design requirements relating to Sankey diagrams. This is true for the specification of design requirements relating to FEW nexus visualizations as well. Because of the lack of design requirement specification in these areas, task taxonomies in graph-based visualization, network evolution, and visual analysis of volume data were instead examined to produce the set of design requirements for the proposed visualization. The basis for these design requirements came significantly from the works of Amar *et al.* (2005), Lee *et al.* (2006), Ahn *et al.* (2014), and Laha *et al.* (2015). This approach was used as it allowed the proposed visualization design requirements to be developed and adapted from visualization design and visual analytics areas that had a much broader range of previous research than was currently available within the area of Sankey diagram design requirements.

The work by Amar *et al.* (2005) focused on low-level visual analytics tasks and identified the ten analytic task taxonomies: retrieve value, filter, compute derived value, find extremum, sort, determine range, characterize distribution, find anomalies, cluster, and correlate. The work by Lee *et al.* (2006) focused on graph-based visual analytics tasks, and identified the four major task categories: topology-based, attribute-based, browsing, and overview. These categories were then further broken into the smaller task categories: adjacency, accessibility, common connection, connectivity, on the nodes, on the links, follow path, and revisit (Lee *et al.*, 2006). In the work by Ahn *et al.* (2014), temporal tasks were the main area of focus for the analysis. This produced a set of higher level tasks which were derived from a set of fifteen real systems and studies. The task categories identified in this work were temporal data processing tasks, individual temporal event tasks, aggregated temporal event tasks, rate of changes, and compound tasks. These tasks were again broken into a number of smaller subtasks, each with their own examples relating to the fifteen works reviewed. The work by Laha *et al.* (2015) focused on the classification of user tasks relating to visual analysis in volume data. In this work, the task categories identified were search, pattern recognition, spatial understanding, quantitative understanding, and shape descriptor. Although many of these tasks do not directly pertain to design requirements relating to Sankey diagrams, they do however present tasks that can easily be adapted for use in Sankey diagram design requirements, or for flow diagram design requirements in general. This is due to Sankey diagrams, and many other flow diagrams, having a number of issues and traits that are similar to those that can typically be found in volume visualization, such as overlaps and visual quantitative estimation.

In this project, four domain experts were consulted in order to gain a better understanding of the tasks and challenges that may be present when trying to develop a visualization design for the FEW nexus. These four domain experts had expertise in sustainability science and international development, hydrology and atmospheric sciences, modeling theories and methodologies, and decision science and sustainability science. Working with these domain experts provided a diverse range of areas of expertise that encapsulated the majority of the disciplines present in the FEW nexus. In the early stages of the visualization design process, domain experts were asked to identify tasks they would like to be able to perform, and questions they would like to be able to answer using a diagram of the FEW nexus. This process started by first asking the domain experts to identify high-level tasks that they may like to be able to perform. These high-level tasks helped to identify areas of the design that may be of significant importance and areas that may not need as much attention, or increased levels of detail. After these high-level tasks were identified, a list of low-level tasks were presented to the domain experts that were felt to be representative of the given high-level tasks. Feedback was then received from the domain experts on whether they felt that these low-level tasks were an accurate representation of their original high-level tasks. This process was then iterated until the domain experts felt that the proposed low-level tasks fit the original high-level tasks that they had posed. This helped to provide a more focused set of design requirements than would otherwise have been possible using only related task taxonomies from other visual analytics areas.

Through the review of the works by Amar *et al.* (2005), Lee *et al.* (2006), Ahn *et al.* (2014), and Laha *et al.* (2015), the previously mentioned discussions with domain experts, as well as the analysis of Water-Energy Sankey diagrams, seven major task categories were identified that applied to the static Sankey diagram designs. Although these task categories were developed specifically with Sankey diagrams in mind, many of these are also applicable to a number of other flow diagram designs and application areas outside of the FEW nexus. The seven task categories that were identified were: object identification, value retrieval, link analysis, path analysis, pattern recognition and column understanding, extremum identification, and inter-sector analysis. Each task category was then further broken into a set of more specific tasks. The sample questions posed in each of the following sections were derived from an examination of the Water-Energy Sankey diagrams produced by Greenberg *et al.* (2017), as well as from further discussions with domain experts.

Before further specifying the proposed task taxonomy, several design elements relating to Sankey diagrams need to first be defined. These terms are based on elements identified within the Water-Energy Sankey diagrams produced by Greenberg *et al.* (2017), some of which do not appear in traditional Sankey diagrams.

Node - A node is an element which presents numerically, as well as based on height, the amount of a resource that is either available to be used or is being used.

Sector - A sector is a set of elements that share a unit of measure, or share some other defining characteristic. In the diagrams produced by Greenberg *et al.* (2017), there are two sectors being presented, the energy sector and the water sector.

Supernode - A supernode is a collection of nodes from a single, or multiple, sectors that are placed in close proximity and share some semantic meaning. This is similar to a supernode in graph theory, as a large number of incident edges are present within supernodes here.

Edge - An edge identifies a connection between nodes and presents the amount of a resource that is flowing between nodes by encoding it in the thickness of the edge.

Column - A column refers to a set of supernodes that share a horizontal position and exhibit a similar semantic meaning. An example from the diagrams produced by Greenberg *et al.* (2017) is the primary resources column, which presents all primary resources that are being used within the diagram.

4.1.1 T1 Object Identification (OI)

In terms of the design requirements for the proposed visualization, object identification refers to identifying specific nodes or supernodes based on a set of given characteristics. The following are examples of object identification tasks, along with sample questions relating to the FEW nexus:

T1.1 Identify existence/absence of a node

  - Is there a Fresh Surface node in the water sector?
  - Is there an Industrial node that is part of the end-use nodes?
  - Does the energy sector have Hydro energy?

T1.2 Identify the name/type of a node

  - What are the names of the consumer nodes?
  - What type of node is the Rejected Energy node?

T1.3 Identify sectors used in a supernode

  - What resource sectors are used in the Residential supernode?
  - Is the energy sector used in the Irrigation supernode?

T1.4 Count the number of supernodes with a given feature

  - How many supernodes use resources from the water sector?
  - How many end-use supernodes use energy?

4.1.2 T2 Value Retrieval (VR)

In terms of the design requirements for the proposed visualization, value retrieval refers to identifying values relating to specific nodes, supernodes, or edges based on a set of given characteristics. The following are examples of value retrieval tasks, along with sample questions relating to the FEW nexus:

T2.1 Amount of inflows/outflows of resources

- What is the inflow of water to the Industrial supernode?
- What is the outflow of energy from the Commercial supernode?
- What is the outflow of Fresh Surface water from the Water Resources supernode?
- What is the inflow of energy to the Rejected Energy node?

T2.2 Amount of resource contributed from a specific parent node or supernode

- What is the amount of water supplied to the Residential supernode from the Fresh Surface node?
- What is the amount of energy contained in the Rejected Energy node that is from the Transportation supernode?
- What is the amount of water contained in the Electricity Generation supernode that is from the Fresh Ground node?

T2.3 Amount of resource produced by a source node

- What is the amount of energy produced in the Biomass node?
- What is the amount of water produced in the Fresh Ground node?

4.1.3 T3 Link Analysis

In terms of the design requirements for the proposed visualization, link analysis refers to identifying high level interactions between adjacent nodes or supernodes of similar and dissimilar types. The following are examples of link analysis tasks, along with sample questions relating to the FEW nexus:

T3.1 Identify parents of a single node or supernode

- Where does the inflow of water in the Residential supernode come from?
- What resources contribute to the generation of electricity?

T3.2 Identify children of a single node or supernode

- What supernodes does the Biomass node energy go into?
- Where does the water outflow from the Residential supernode go?

T3.3 Identify splits/merges of different resource types

- After being processed (or consumed) in the Commercial supernode, what forms are the energy transformed into? Do they go into the same or different destinations?
- After being processed (or consumed) in the Electricity supernode, what forms are the water transformed into? Do they go into the same or different destinations?

T3.4 Identify conversions from one resource into another resource

- What resources are used to generate electricity?

4.1.4 T4 Path Analysis (PA)

In terms of the design requirements for the proposed visualization, path analysis refers to identifying high level interactions between nodes and edges from the source to the final destination. The following are examples of path analysis tasks, along with sample questions relating to the FEW nexus:

T4.1 Identify existence/absence of a path

  - Does any water from Fresh Surface water node go to the Surface Discharge node?
  - Does any energy supplied by the Biomass energy node go to the Rejected Energy node?

T4.2 Find a path of specific resource from the source to a potential final destination

  - What supernodes are encountered along the path from the Fresh Surface water node to the Surface Discharge node?
  - What supernodes are encountered along the path from the Biomass energy node to the Rejected Energy node?
  - How much Fresh Surface water goes to the Surface Discharge node?

T4.3 Find all paths of specific resource from the final destination to the potential sources

  - For all the Rejected Energy, where did each inflow come from? Which processors or consumers did they pass?
  - For all the Surface Discharge, where did each inflow come from? Which processors or consumers did they pass?

4.1.5 T5 Pattern Recognition and Column Understanding (PRCU)

In terms of the design requirements for the proposed visualization, pattern recognition and column understanding refers to the identification of trends or repeated characteristics occurring within a single or multiple diagrams. The following are examples of pattern recognition and understanding tasks, along with sample questions relating to the FEW nexus:

T5.1 Intra-column understanding

- Which end use supernodes use similar water resources?
- What energy resources are used in at most one end use supernode?
- Are spikes/drops in a resource's usage between the years 1990 and 2010 similar to spikes/drops of other resources in the same year? If so, in which years does this occur?
- What is the ordering, in descending order, of Petroleum usage in the end use supernodes?

T5.2 Patterns

- Considering only the growth rate for Natural Gas usage between the years 1990 and 2010, is it likely Natural Gas usage increased between the years 2010 and 2011?
- How did the distribution of Fresh Surface water change between the years 1990 and 2010?
- Did any droughts occur between the years 1980 and 2005 with similar characteristics in the water sector to the characteristics experienced in the drought in 2012?

4.1.6 T6 Extremum Identification (EI)

In terms of the design requirements for the proposed visualization, extremum identification refers to identifying the extreme values either locally or globally within the visualization. The following are examples of extremum identification tasks, along with sample questions relating to the FEW nexus:

T6.1 Identify maximum/minimum local resource type supply/usage

- Which end-use supernode used the largest amount of Fresh Surface water in the year 2010?
- Which primary energy source provided the smallest amount of energy in the year 2010?
- Which end-use supernode contributed the largest amount of electricity to the Rejected Energy node in 2010?

T6.2 Identify maximum/minimum global resource type supply/usage

- What is the minimum amount of water consumed by an end-use supernode and in which year?
- Which primary energy source contributed the most energy in a single year to the Electricity Generation node and in which year?
- Which primary energy sources contributed the most total energy to the Electricity Generation node?
- Do the minimum amount of water consumed and the maximum amount of water consumed by individual end-use supernode occur within the same year?

4.1.7 T7 Inter-sector Analysis (IA)

In terms of the design requirements for the proposed visualization, inter-sector analysis refers to the analysis of inter-sector influence and interaction within a single or multiple diagrams. The following are examples of inter-sector analysis tasks, along with sample questions relating to the FEW nexus:

T7.1 Inter-sector influence

- Which sector, the energy sector or the food sector, relies more heavily on the Fresh Ground water node in terms of usage between the years 2010 and 2015?
- Does the growth rate of the Fresh Surface water node impact the growth rate of the food sector?
- During which year does electricity generation limit Fresh Surface water usage?
- Has increased water usage by the energy sector impacted usage by the food sector?

T7.2 Inter-sector resource usage

- What is the total amount of energy that is used in the water sector between the years 2008 and 2013?
- What is the total amount of water used to produce biomass between the years 2010 and 2012?
- What is the total amount water that was consumed or evaporated from use by the energy sector between the years 2008 and 2010?

4.2 Initial Network Embodied Sectoral Trajectory (NEST) Approaches

Visualization of the FEW nexus is a difficult task that heavily relies on the framing of the problem for the analysis that is intended. For this reason, a number of potential visualization approaches were proposed and examined. Each of these offered solutions to some of the current issues in FEW visualization but also posed their own limitations and issues. The proposed visualization approaches, along with justification for the design and the primary issues encountered, can be seen in the following sections.

4.2.1 Hierarchical Graph-Based Visualization

The first proposed visualization approach was a Hierarchical Graph-Based Visualization. This method was proposed for 2 key reasons. First, when viewing current FEW research, similar diagrams of the FEW nexus are often used as a way to present the interactions among the sectors at a high level. Second, although there are interactions across sectors within the FEW nexus, each of the sectors inherently contain the other sectors. A sample of the structure of this visualization approach can be seen in Figure 4.1. In this visualization, the sizing of parent nodes are based on the sum of the usage of its child nodes. Focusing on just the Water sector in this visualization, it can be seen that the Food sector appears in the form of Agriculture and Urban Agriculture, and the Energy sector appears as Hydropower and Cooling. Because the external sectors are represented within the hierarchy of the supplying sector, this allows a high level description of the nexus to be presented.

With this kind of approach, there are a number of significant limitations that appear. The major limitation that appears, and is an issue with a number of FEW visualizations, is the issue of scale. Approaching the nexus in this way presents each of the sectors as a single entity rather than a number of smaller components. While

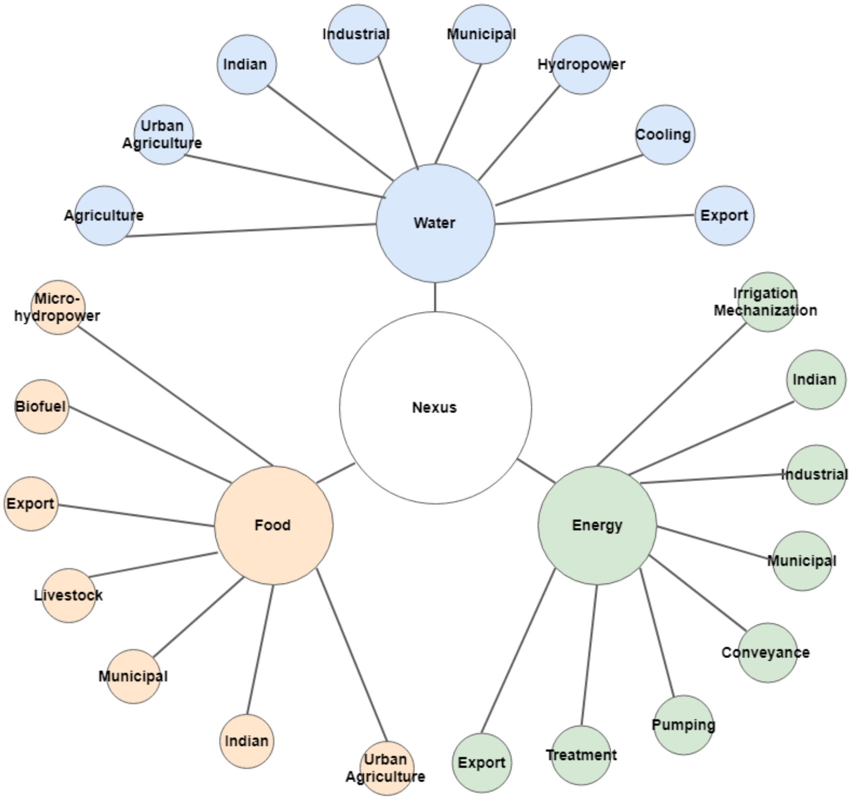


Figure 4.1: FEW hierarchical graph-based visualization structure. The center node (Nexus) is the root, with child nodes Food, Energy, and Water. The radii of the children of a node sum to the radius of the node.

at a national, or potentially statewide, scale this is likely not an issue, viewing the nexus in this way at a city or regional scale could be problematic, and highly abstracts the sources that make up each of the sectors. The other major issue with this visualization approach is the difference in units of measurement across the sectors. In an attempt to circumvent this issue, the sizing of the nodes was based on a dollar value rather than the original unit of each of the sectors. Although a potentially reasonable compromise in a number of situations, in terms of the FEW nexus this removed the ability to compare usage within a single sector. This is due to the fact that pricing within the subcategories for each sector can vary based on the intended recipient, for example, municipal water use versus agricultural water use. The final major limitation is that reuse and waste of resources are not able to be easily represented in a visualization with this kind of structure. Without including reuse and waste, more abstraction is added, this time in terms of the flow of the resources through the nexus. Although again, at a larger scale, this is likely not an issue, at smaller scales this information could be significantly important.

4.2.2 Treemap Motivated Visualization

The next proposed visualization approach was a Treemap Motivated Visualization. The key reasons for proposing this visualization approach were that flow could be represented through the nexus, from supply to output/waste, and a reduction of the negative impact of difficult to follow edge crossings that can occur from complex interactions could be achieved. Sample structures of this visualization approach can be seen in Figure 4.2 and Figure 4.3. In these figures, the height of the boxes represent the amount of the resources used, in the unit of the respective sector. As these figures are used to only represent the structure of the visualization, they do not display the amount of the resource that was used, but these values would be displayed in an application using this approach. In Figure 4.3, edges were introduced as visual markers to help locate where a resource was going to or coming from. These edges were used only to signify a connection between boxes, and do not represent the amount of resources being used. By using uniform sized edges as visual markers only, the issue of small edges is eliminated and the impact of edges that crossover each other should be slightly reduced since the edges serve as only a visual marker to help identify the important information.

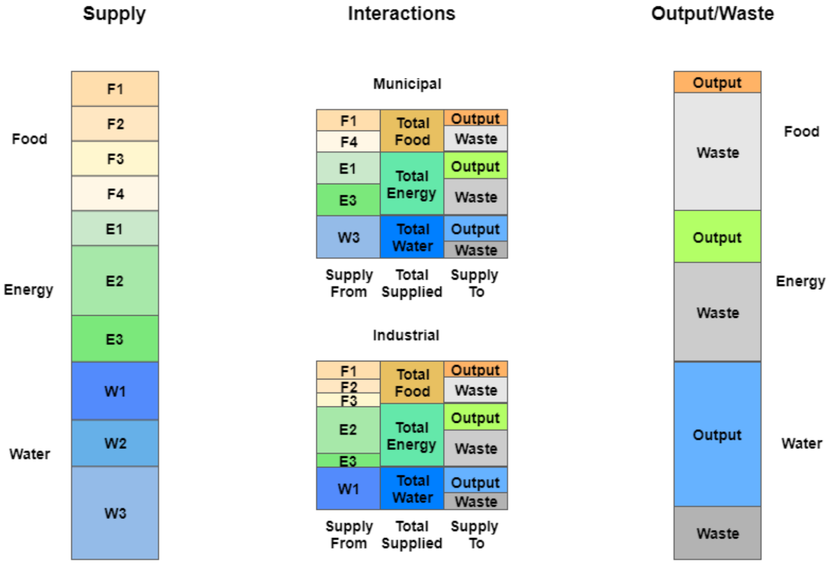


Figure 4.2: Treemap motivated visualization structure without connecting edges. The Supply column presents all resources available in the example area, the Interactions column presents the amount of resources that go into and come out of each end-use area (i.e. Municipal and Industrial), and the Output/Waste column presents whether resources are reusable or wasted/consumed.

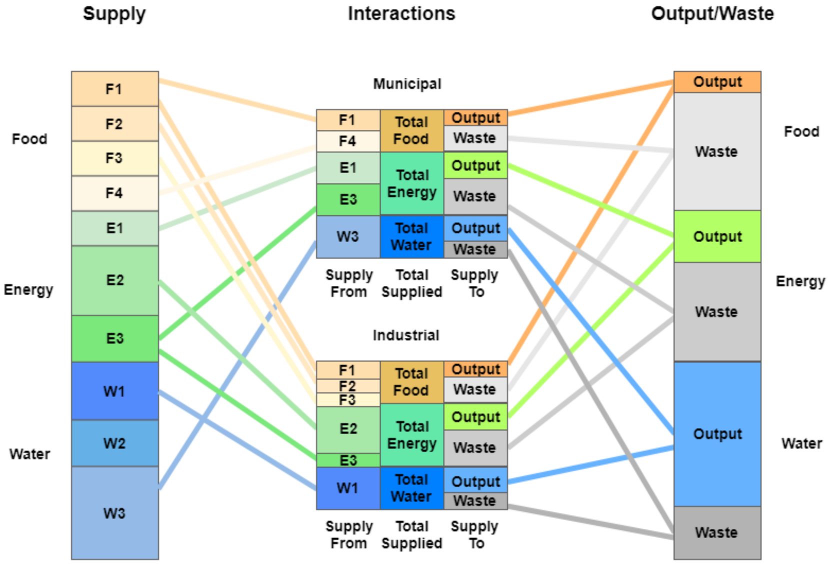


Figure 4.3: Treemap motivated visualization structure with connecting edges. The Supply column presents all resources available in the example area, the Interactions column presents the amount of resources that go into and come out of each end-use area (i.e. Municipal and Industrial), and the Output/Waste column presents whether resources are reusable or wasted/consumed. Edges are used to aid in identifying where a resource is flowing to or coming from.

This visualization approach comes with some limitations and issues. One large issue is the reliance on color to distinguish different sectors, subcategories, and other similar structures. Without using a similar color scheme within a sector, identifying the subcategories in the sector can be difficult and may potentially add increased search time. However, maintaining a consistent color scheme can become significantly problematic if a sector is made up of many subcategories, as only a finite number of distinguishable similar colors are available. The other important issue with this approach is the potential for difficulty understanding the visualization. Although this visualization generally is not difficult to understand once an explanation of the flow has been given, on initial inspection the visualization may be unclear and may make this unrealistic for use in static presentations. The key limitation is that this type of visualization is not highly suitable for comparison across similar visualizations with different values. This is because using a constant scaling value for all of the visualizations may result in either significantly large boxes that make up most of the visualization, or very small boxes that are indiscernible. In terms of larger scales, this is likely not an issue as the small boxes do not represent a significant fraction of a resource, but in terms of smaller scales these could be important resources.

4.2.3 Road Network Motivated Visualization

The final proposed visualization approach was a Road Network Motivated Visualization. The key reasons for proposing this visualization approach were that flow could be represented through the nexus, the negative impact of edge crossings could be reduced, and the underlying structure is reminiscent of a road network grid. A sample structure of this visualization approach can be seen in Figure 4.4. The use of a road network grid approach allows for a visualization that, although not traditionally used, should be understandable for most individuals, at least in a general sense. In this visualization approach, the thickness of the edges and the height of the boxes represent the amount of resource being used. The structure representation in Figure 4.4, like the Treemap Motivated Visualization structure representation, does not display the resource values, but they would be displayed in an application using this approach. By using this approach, although there is still crossover of edges, the visual impact of edge crossings should be reduced as they are now perpendicular and still allow a substantial portion of the vertical boxes to be seen.

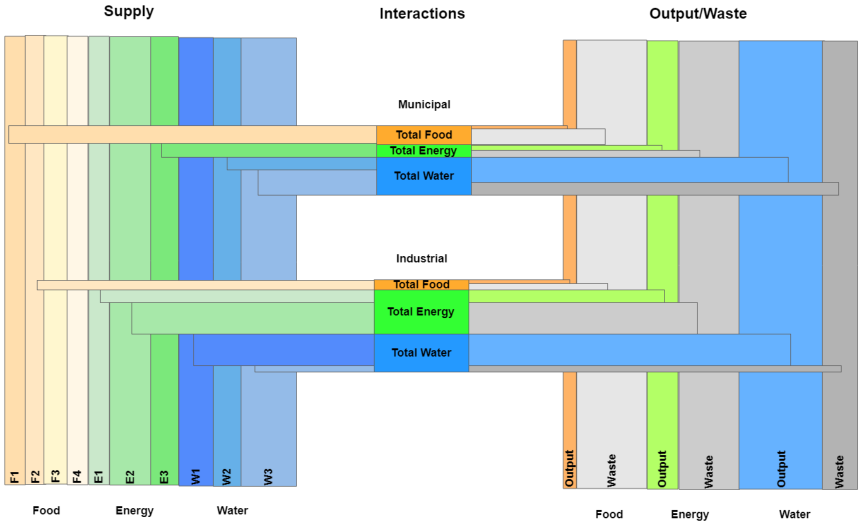


Figure 4.4: Road network motivated visualization structure. The Supply column presents all resources available in the example area, the Interactions column presents the amount of resources that go into and come out of each end-use area (i.e. Municipal and Industrial), and the Output/Waste column presents whether resources are reusable or wasted/consumed. Edges are used to represent amount of resource flowing into or coming from a node.

Again, this visualization approach comes with limitations and issues. This visualization approach, like the Treemap Motivated Visualization, suffers from the issues of reliance on color to distinguish sectors, subcategories, and other similar structures, finite distinguishable similar colors, and issue of scale. The key issue in this visualization however is the use of perpendicular edges. As the horizontal edges in the center of the visualization make up a portion of either the vertical edges on the left or right of the visualization, attempting to determine how large of a portion could be difficult visually since they are not in the same direction. Because of this, it could potentially be difficult to visually determine whether a source is being entirely used. In terms of limitations, this again suffers from potentially being unclear when first presented and is not suitable for comparison across similar visualizations with different values.

4.3 Iterative NEST Designs

After much consideration, the Treemap Motivated design from Section 4.2.2 was selected as the basis for the remainder of the visualization design process. Although this initial design suffered from a number of issues, many of which were mitigated during the iterative visualization designing, it appeared to best fit the design requirements from Section 4.1, and also appealed the most to a small sample of domain experts when compared to a number of other potential designs. This approach also contained characteristics that provided a simpler and more intuitive way for domain experts, as well as non-experts, to explore the data using the interactive visualization than the other proposed designs. The iterative visualization design stage was defined by 4 major design iterations which are described in detail in the following sections.

4.3.1 NEST Iteration 1

The first major design iteration introduced inter-sector interactions between the different sectors of the FEW nexus. This was done by adding Sectors and Intermediate Resources columns to the original diagram. A sample structure of this iteration of the visualization approach can be seen in Figure 4.5.

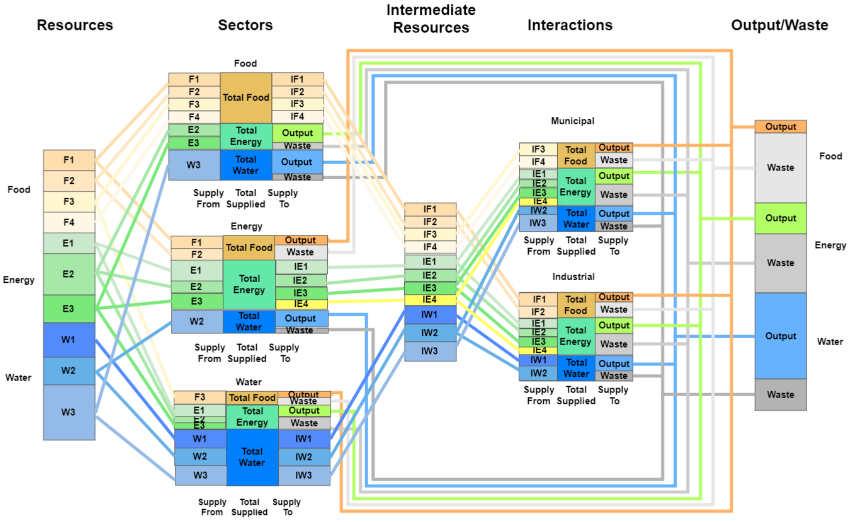


Figure 4.5: NEST iteration 1 structure, representing a sample FEW system. The Resources column presents available resources in this sample system. The Sectors column presents the resources actually being used in the system, and includes use of resources in other sectors (i.e. W2 being used in the Energy sector). The Intermediate Resources column is used to recombine resources into a common area of the diagram. The Interactions column presents the amount of resources that go into and come out of each end-use area (i.e. Municipal and Industrial). The Output/Waste column presents whether resources are reusable or wasted/consumed. Edges around the outside of the diagram are used to route wasted resources from intermediate columns of the diagram to the Output/Waste column without interfering with the rest of the diagram.

This, however, came with a large number of complex interactions, as well as a large number of difficult to manage edge crossings. This also lead to the identification of several new issues which were:

- Does it make sense to abstract what resources are specifically used for in other resource sectors of the static diagram?
- How do we represent amount of Supply From that occurs in the Supply To?
- Are conversions from one resource to another resource always equal to each other (i.e. Does the amount of energy required to generate electricity equal the amount of energy produced by electricity generation)?

- How do we add water treatment and electricity generation in a strategic way?
- How do we better layout Output/Waste so that edges aren't extending the majority of the diagram?
- Can we reorganize some of the nodes within supernodes to reduce the edge crossings, while still maintaining the integrity of the diagram?

4.3.2 NEST Iteration 2

The second major design iteration introduced bundled edges, primary to secondary resource conversion and wastewater treatment, and a number of design elements utilized in the Road Network Motivated design from Section 4.2.3. The major change to this diagram was the decision to split the Output/Waste column into two pieces, one above and one below the rest of the diagram respectively. This was done to minimize the number of edges that extended the majority of the diagram, horizontally and vertically, and reduce clutter around the Wastewater Treatment supernode. Doing this also provided room for the Electricity Generation supernode and the Wastewater Treatment supernode without significantly widening the overall diagram. The addition of bundled edges allowed the number of edge crossings to be significantly reduced without much abstraction. A bundled edge was used when three or more edges from a source node sector within a supernode connect to a single target supernode. A sample structure of this iteration of the visualization approach can be seen in Figure 4.6.

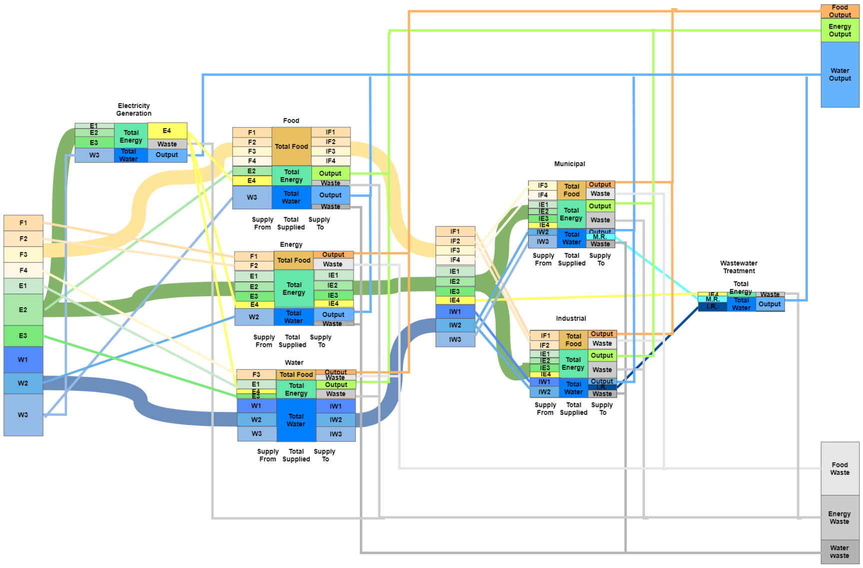


Figure 4.6: NEST iteration 2 structure, representing a sample FEW system. Similar columns are used as in Nest iteration 1. Electricity Generation and Wastewater Treatment are added to represent conversions from primary resources to secondary resources and wastewater to usable water respectively. Edge bundling is also introduces when 3 or more edges go from one set of similar resource sector nodes to the same location. The Output/Waste column is separated to allow resources to flow around the Wastewater Treatment.

Although this design did provide solutions for some of the questions in Section 4.3.1, and some questions were accepted as issues in Sankey diagrams as well, three key questions still remained:

- Does it make sense to abstract what resources are specifically used for in other resource sectors of the static diagram?
- How do we add water treatment and electricity generation in a strategic way?
- Can we reorganize some of the nodes within supernodes to reduce the edge crossings, while still maintaining the integrity of the diagram?

4.3.3 NEST Iteration 3

The third major design iteration mainly focused on optimizing the layout of the supernodes from the previous iteration. The major change in this iteration was the decision to move the Electricity Generation and Wastewater Treatment supernodes to the upper part of the diagram. This allowed supernodes that represent conversions from one resource into another resource to be contained in a row at the top of the diagram, rather than in the middle. This also allowed the Output and Waste supernodes to be moved back towards the middle, which reduced the distance that edges were traveling from the end use column. Another major design change was to separate the set of resources in the first column based on their sector. This was done to help identify the colors for each of the resource sectors and to allow more spaces for the edges coming from these supernodes. The final major change was that the intermediate nodes column was removed. This was done because the removal of the column did not appear to decrease readability of the diagram but did reduce the overall width. A sample structure of this iteration of the visualization approach can be seen in Figure 4.7.

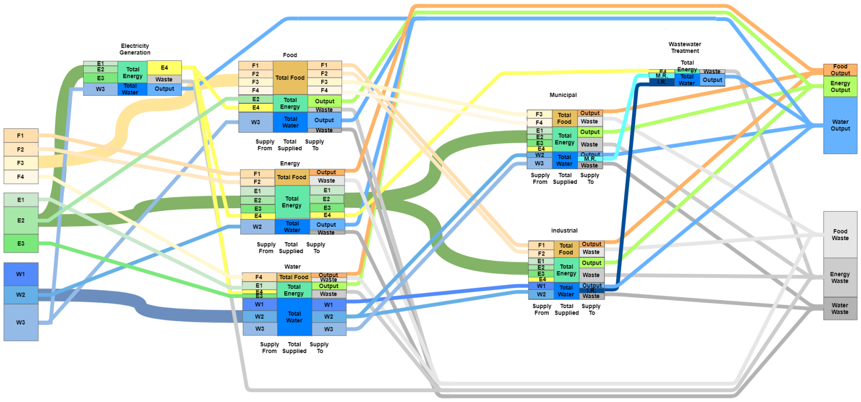


Figure 4.7: NEST iteration 3 structure, representing a sample FEW system. Similar elements are used as in Nest iteration 2. Electricity Generation and Wastewater Treatment have been moved to a common row to present that they are conversions of resources. The Output/Waste column are moved back toward the center of the design, vertically, as that space is no longer occupied by Wastewater Treatment. The Intermediate resources column has been removed to eliminate unnecessary abstraction.

This design iteration did, however, leave two of the questions from Section 4.3.1 unanswered. This was done as these questions were highly dependent on end user preference and not optimization of the overall placement of a supernode. Thus, these questions were left for feedback from domain experts, as well as other potential end users, as changes to these characteristics could make the visualization potentially difficult to understand. The two questions that remained after this design iteration were:

- Does it make sense to abstract what resources are specifically used for in other resource sectors of the static diagram?
- Can we reorganize some of the nodes within supernodes to reduce the edge crossings, while still maintaining the integrity of the diagram?

4.3.4 NEST Iteration 4

The final major design iteration focused on how to handle small nodes. This became a significant issue because a large number of the nodes could become quite small and may not actually be able to be drawn. This issue is often mitigated in Sankey diagrams by setting a minimum size, but in the proposed visualization this could not have been done without losing some of the diagrams readability. The solution that was chosen was to first shrink the width of the current nodes and remove their labels. Doing this gave a quantitative estimation representation of the data, which appeared to be the strong point of traditional Sankey diagrams. Then columns were added to the left and right of each supernode, excluding the source and destination supernodes which only contained a right column and left column respectively, where the left columns included all of the inflows and the right columns included all of the outflows. This was done to present the numeric value and label of each left column and right column node from the qualitative representation at either a uniform or scaled size. Uniform sizes were used to only help present the connection between two nodes, while also minimizing the impact of the new nodes on the quantitative estimation understanding of the entire diagram. The scaled sizes were used to help present the connection between two nodes while also maintaining the size distribution within a column, by scaling the size of the nodes based on the smallest node size. The negative of the scaled approach is the potential impact on the quantitative estimation in the entire diagram, as a set of scaled connection nodes cannot be compared to the scaled connection nodes from any other supernode or to the other set of connection nodes within the same supernode. A sample structure of this iteration of the visualization approach can be seen in Figure 4.8.

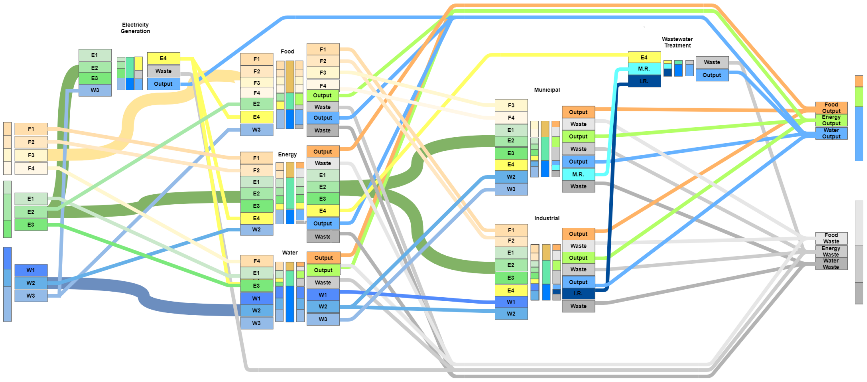


Figure 4.8: NEST iteration 4 structure, representing a sample FEW system. Similar elements and positions are used as in Nest iteration 3. Internal nodes have been added to represent the use of a resource based on the size of the node and previous nodes are now used as connection points for edges and to show the name and value associated with the new nodes.

4.4 Data Structure

After the iterative designing was complete, a data structure was needed to store this new design. The structure decided on was similar to the data structure for a Sankey diagram where only nodes and edges are defined, just with a number of added attributes in the nodes. In the input data, edges were characterized only by the source, target, and value, which is the structure generally used for Sankey diagrams. The value contained in the edge was used to determine the size of the nodes, as the edges were a uniform size. In the input data, the nodes were characterized by name, ID, cluster, type, sector, units, and depth. The other key difference between this structure and a general Sankey diagram data structure is that edges can go from a strictly output node or into a strictly input node at any point in the diagram, not only the endpoints. This is done, because in some supernodes, a resource can be changed into some other resource. Due to this, some nodes appear as near duplicates where one node represents input of a resource to a supernode and a node with similar values represents the output of the resource from the supernode. A sample JSON file of this structure can be seen in Figure 4.9.

```
{
"Nodes": [
    {"name":"Energy 0", "id":"Resources_Energy0_O","cluster":"Resources", "type": "Output","sector":"Energy","units":"Trillions of BTU/Year","depth":0},
    {"name":"Energy 1", "id":"Resources_Energy1_O","cluster":"Resources", "type": "Output","sector":"Energy","units":"Trillions of BTU/Year","depth":0},
    {"name":"Water 0", "id":"Resources_Water0_O","cluster":"Resources", "type": "Output","sector":"Water","units":"Millions of Gallons/Day","depth":0},
    {"name":"Water 1", "id":"Resources_Water1_O","cluster":"Resources", "type": "Output","sector":"Water","units":"Millions of Gallons/Day","depth":0},

    {"name":"Energy 0", "id":"EndUse_Energy0_I/O","cluster":"End Use", "type": "Input/Output","sector":"Energy","units":"Trillions of BTU/Year","depth":1},
    {"name":"Energy 1", "id":"EndUse_Energy1_I/O","cluster":"End Use", "type": "Input/Output","sector":"Energy","units":"Trillions of BTU/Year","depth":1},
    {"name":"Water 0", "id":"EndUse_Water0_I/O","cluster":"End Use", "type": "Input/Output","sector":"Water","units":"Millions of Gallons/Day","depth":1},
    {"name":"Water 1", "id":"EndUse_Water1_I/O","cluster":"End Use", "type": "Input/Output","sector":"Water","units":"Millions of Gallons/Day","depth":1},

    {"name":"Energy 0", "id":"Output_Energy0_I","cluster":"Output", "type": "Input","sector":"Energy","units":"Trillions of BTU/Year","depth":2},
    {"name":"Energy 1", "id":"Output_Energy1_I","cluster":"Output", "type": "Input","sector":"Energy","units":"Trillions of BTU/Year","depth":2},
    {"name":"Water 0", "id":"Output_Water0_I","cluster":"Output", "type": "Input","sector":"Water","units":"Millions of Gallons/Day","depth":2},
    {"name":"Water 1", "id":"Output_Water1_I","cluster":"Output", "type": "Input","sector":"Water","units":"Millions of Gallons/Day","depth":2},
    ],

 "Links":[
     {"source":"Resources_Energy0_O","target":"EndUse_Energy0_I/O","value":10},
     {"source":"Resources_Energy1_O","target":"EndUse_Energy1_I/O","value":15},
     {"source":"Resources_Water0_O","target":"EndUse_Water0_I/O","value":10},
     {"source":"Resources_Water1_O","target":"EndUse_Water1_I/O","value":20 },

     {"source":"EndUse_Energy0_I/O","target":"Output_Energy0_I","value":10},
     {"source":"EndUse_Energy1_I/O","target":"Output_Energy1_I","value":15},
     {"source":"EndUse_Water0_I/O","target":"Output_Water0_I","value":10},
     {"source":"EndUse_Water1_I/O","target":"Output_Water1_I","value":20},

    ]
 }
```

Figure 4.9: Sample JSON data file showing the structure of the data used to generate a NEST diagram.

The name attribute was used to specify the label for the node, as well as to identify the same resource appearing multiple times for coloring purposes. The id was a unique value for each node and was used to identify a specific node. The cluster attribute was used to identify what cluster, or supernode, a node was part of to allow the nodes of a similar cluster to be positioned together. The type attribute was used to identify whether a node was input, output, or input/output. This was used to determine how the node would be drawn and how edges could be attached to it. The sector attribute was used to specify the different categories of resource sectors that were available. In terms of Figure 4.8, there were 3 sectors: Food, Energy, and

Water. This was done to scale the size of the nodes based on the sector that they belonged to and for coloring purposes. The next attribute, units, was used to specify the unit of measurement that the node was in. Finally, the depth attribute was used to specify the column that a node would be placed in. This was done because although the diagram could be automatically generated without it, the semantic meaning of a node may not match where it was placed. This can be seen in the simple example represented in Figure 4.10. On side a, the nodes are placed in the column as far to the left as they can be. On side b, the nodes are placed in the column as far to the right as they can be. Although this appears to be a trivial problem, the issue is whether node B is most semantically similar to node A or node C. If column one represents resource nodes, column two represents end-use nodes, and column 3 represents final destination nodes, the placement of node B could be significantly important. This implied meaning is not able to be represented by just following a set of rules, it is an attribute that has to be specified in the data itself.

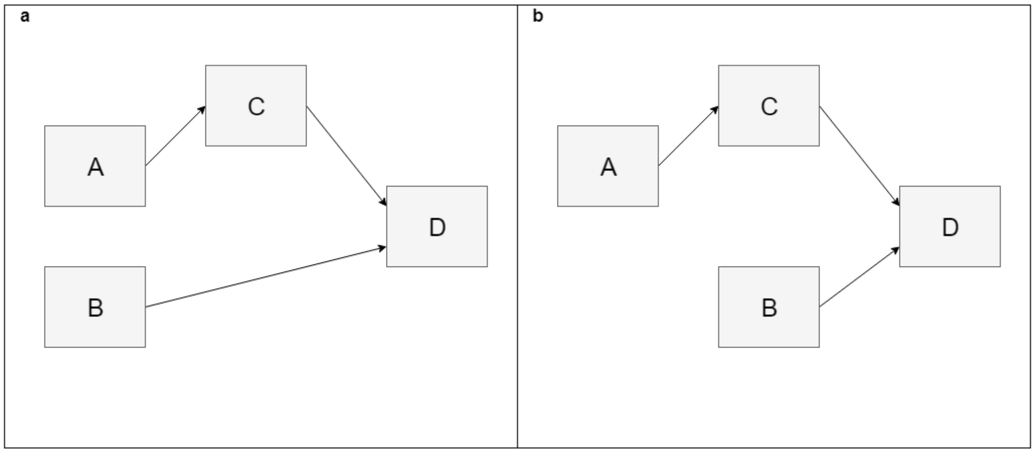


Figure 4.10: Sample column selection problem highlighting the issue of selecting where node B should be placed in relation to the other nodes.

The final visualization design, henceforth referred to as the Network Embodied Sectoral Trajectory (NEST) diagram, is a set of visualization design alternatives that incorporate characteristics from Sankey diagrams, treemaps, and graphs, to improve readability and minimize the negative impact of edge crossings on the understandability of the diagram. At this stage, it was decided that edge bundling would not occur in the static diagrams because it may be considered confusing, but would instead be used for the interactive diagrams. The NEST design alternatives were implemented using JavaScript, as well as the D3 library developed by Bostock *et al.* (2011). JavaScript was chosen as it would allow the visualizations to be usable in web applications, thus increasing the potential for collaboration among a larger number of individuals than with a desktop application. A sample of the NEST diagram representing an energy-water system can be seen in Figure 4.11.

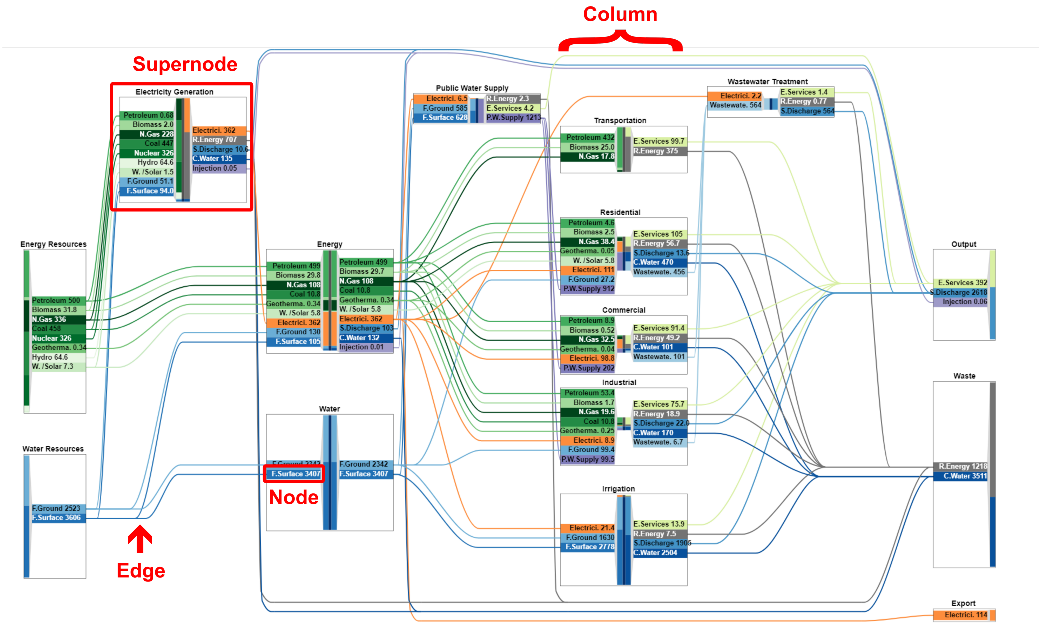


Figure 4.11: Sample NEST diagram design representing a state-level energy-water system. This sample consists of 7 columns containing 15 supernodes that are connected by edges.

4.5.1 Design Elements

For the proposed designs, the 5 major design elements previously identified were expanded upon: nodes, sectors, supernodes, edges, and columns, each of which will be further discussed in the following sections. Although different definitions were needed for the specification of some of these elements, many of them share similar meanings or characteristics to their corresponding design element in Sankey diagrams.

Nodes

A node in the proposed designs refers to an element that presents the amount of a resource either numerically or as a quantitative estimation based on height. This is generally similar to nodes in Sankey diagrams. However, in these designs 5 different categories of nodes are used, which present the majority of the data in the proposed diagrams. These categories, seen in the supernode highlighted in Figure 4.11, from left to right are input connection, input, total, output, and output connection. Input connection and output connection are the nodes that edges are connected to and provide a numeric value for the node as well as the name of the node that a node is connected to. Input, output, and total nodes are used to present a quantitative estimation of the nodes and can be compared to other similar nodes in the diagram by examining their height, which presents the amount of usage. These nodes draw inspiration from treemaps, in that, the total amount of resources used in a supernode is subdivided into the resources that contribute to or are produced in that supernode. This approach is also similar to how edges function in traditional Sankey diagrams, without requiring large edges to span large areas of the diagram. Looking at the top supernode in the middle column of Figure 4.12-1, the leftmost two nodes are input connections nodes showing that 40 units of an energy resource and 75 units of a water resource are coming into this supernode. The two nodes to the right of these are input nodes and represent these values by the height of the nodes, the energy node is 2/3 the height of the node with 60 units contributing to it. The nodes to the right of these nodes are total nodes and provide the sum of the resource sectors. As there is only one energy node and one water node, the height of these nodes is the same as the input nodes. The nodes to the right of these are output nodes, which function similarly to input nodes. These present how much of each of the total nodes is going to where next by the height of the node. Finally, the last nodes are output connection nodes and provide a numeric representation of the output nodes, showing that 40 units of an energy resource is going to the energy resource in the final column and 75 units of a water resource are going to the water resource in the final column.

Sectors

A sector in the proposed designs refers to a set of elements that contain a shared unit of measure, or other defining characteristic. Elements of different sectors do not explicitly interact with each other. This is similar to sectors in Sankey diagrams.

Supernodes

A supernode in the proposed designs refers to a collection of nodes from a single, or multiple, sectors that share a semantic meaning and are placed in close proximity. This is similar to a supernode in the Sankey diagram, with a more diverse set of nodes being contained within the supernode.

Edges

An edge in the proposed designs, like in Sankey diagrams, refers to a connection between 2 nodes. The major difference being that in Sankey diagrams, the thickness of the edge also represents the value of a resource that is flowing through it, whereas in the proposed designs uniform edge sizes are maintained, much like a graph. The exception to this being bundled edges, which appear only in the interactive design, that are thicker than standard edges as they represent a connection between multiple nodes from the same sector in a supernode and multiple nodes from the same sector in another supernode.

Columns

A column in the proposed designs refers to a set of supernodes that share a horizontal position and exhibit a similar semantic meaning. This is similar to columns in Sankey diagrams.

4.5.2 NEST Designs

Three different Nest diagram design alternatives were developed in total. While overall very similar, small characteristic changes were introduced in an attempt to mitigate different issues that may be experienced from the different design choices. A sample representation of a hybrid energy-water system using the three different design alternatives is shown in Figure 4.12.

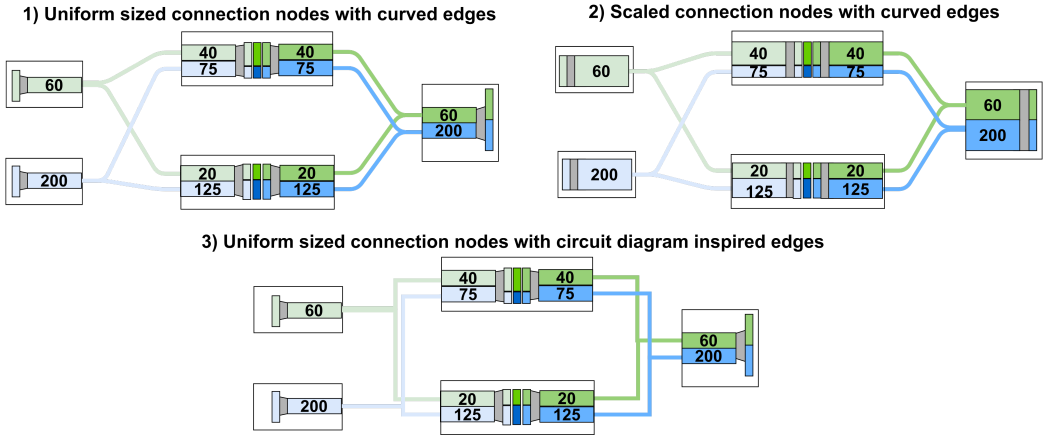


Figure 4.12: Samples of the three NEST diagram design alternatives

The first design proposed was a design using uniform sized connection nodes with curved edges. Uniform sizes were used to only help present the connection between two nodes, while also minimizing the impact of the new nodes on the quantitative estimation understanding of the entire diagram. The negative of the uniform approach being that all connection nodes are represented as the same size and could cause confusion since the input, output, and total nodes are given sizes based on their value. A sample of this design can be seen in Figure 4.12-1.

The second design proposed was a design using scaled connection nodes with curved edges. The scaled sizes were used to help present the connection between two nodes while also maintaining the size distribution of input connection or output connection nodes within a supernode, by scaling the size of the nodes based on the smallest node size. The negative of the scaled approach being the potential impact on the quantitative estimation in the entire diagram as a set of scaled connection nodes cannot be compared to the scaled connection nodes from any other supernode or to the other set of connection nodes within the same supernode. This approach could also potentially yield a much taller diagram, as any supernode that contains a node that is not at least the minimum size, is increased by the minimum size multiplied by the number of either input or output connection nodes in the set containing the node that is not the minimum size. A sample of this design can be seen in Figure 4.12-2.

The final design proposed was a design using uniform sized connection nodes with circuit diagram inspired edges. The circuit diagram approach was used to provide straighter edges that would use less of the space in the diagram and eliminate having to try to follow curved edges. The negative of the circuit diagram approach is that the edges are significantly close together and tracing these edges may be overwhelming without previous experience in circuit diagrams. A sample of this design can be seen in Figure 4.12-3.

The general design for these largely benefited from the idea of artificial nodes proposed by Alemasoom *et al.* (2016). In their work, artificial nodes were used to avoid edge and node intersections within a Sankey diagram. This idea was used in the NEST diagrams to route edges from intermediate columns of the diagram to the Output/Waste column of the diagram without conflicting with the proceeding columns. A sample of this can be seen in Figure 4.13, where the black boxes represent the generated artificial nodes.

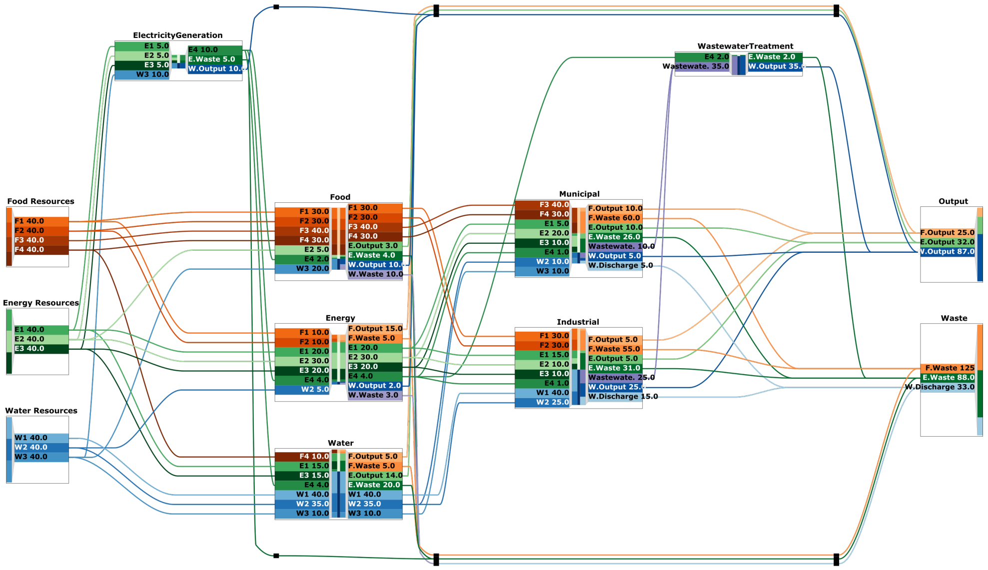


Figure 4.13: NEST diagram with artificial nodes shown as black boxes.

4.6 Interactive NEST Visualizations

The structure of the interactive visualizations are similar to that of the static NEST visualizations, with increased exploration capability. One key interaction focus was value retrieval. In the static NEST visualizations, as well as static Sankey diagrams, value retrieval can be a potentially difficult task when the user is only provided the different height of nodes or edges to determine the value of a node or edge. This means that the initial resource nodes, or larger nodes that can fit text, often have to be used to determine how much of a resource is used at another point in the diagram. Even with the different heights, the value determined by the user will likely not be exactly the value contained in the edge or node. With the addition of interaction, values can be represented for all nodes and edges by hovering over the node or edge, which allows the exact value, even for small nodes, to be viewed quickly and easily.

Another key interaction focus was highlighting. This was done to increase the ability to distinguish different sectors as well as resource flows. By hovering over a node or edge, the nodes or edges that were not related to it became mostly transparent, allowing for the resource path to easily be identified. This was most beneficial when locating where small amounts of a resource were going without having to examine a large portion of the diagram carefully. A sample of this can be seen in Figure 4.14.

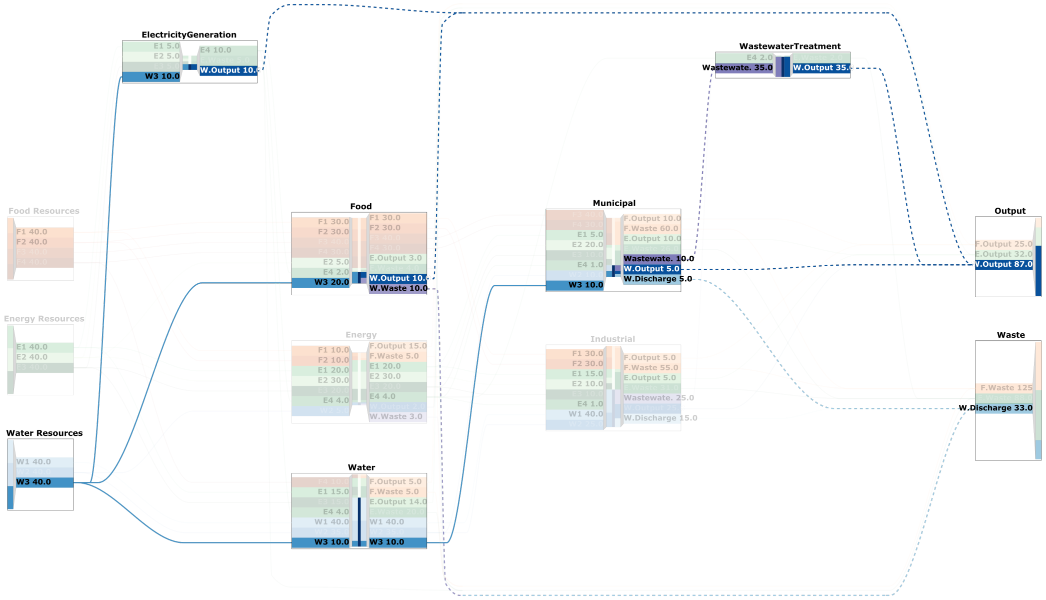


Figure 4.14: Highlighting W3 Node. Solid edges represent known resource flow, and dashed edges represent potential flow as a definite source to final destination flow is often not possible

The final interaction focus was edge bundling. The addition of bundled edges allowed the number of edge crossings to be significantly reduced without much abstraction. In this case, edge bundling refers to the merging of edges from a single sector of a source supernode connecting to the same target supernode. Edge bundling was allowed for two or more edges, as bundling with only two edges presented a near graph representation of the supernodes in Figure 4.15.

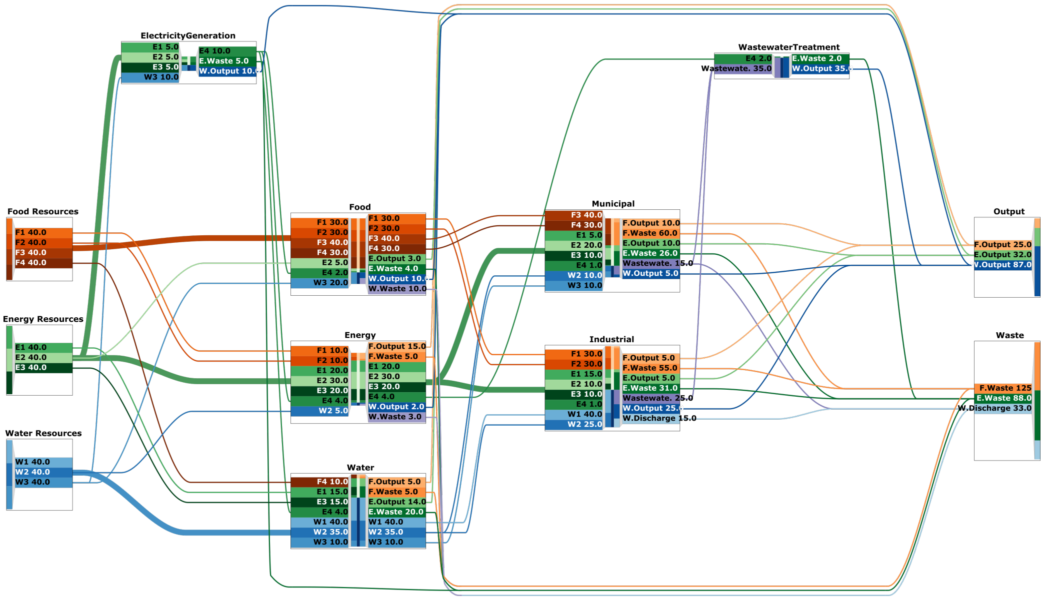


Figure 4.15: Sample NEST diagram with edge bundling used when three or more edges from the same sector are connecting two supernodes

Chapter 5

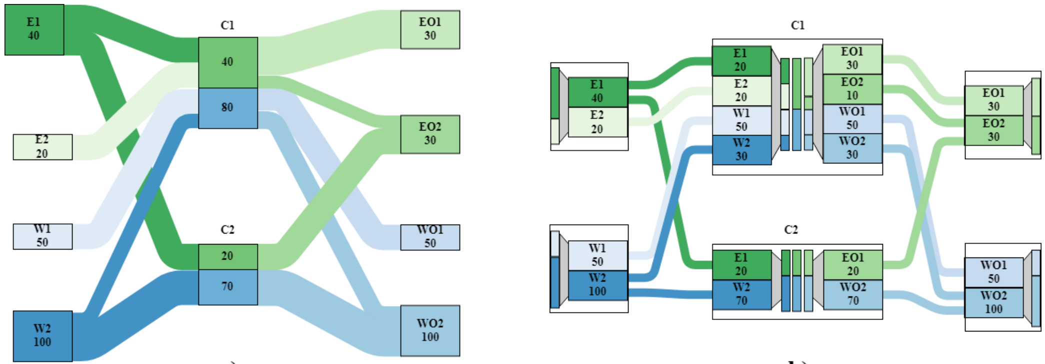
EVALUATION

5.1 Expert Review

After finalizing the NEST diagram designs, the domain experts were presented the diagram designs as a pilot study with a small subsample of the questions that were to appear in the user study. This was done to get feedback on the diagram designs and to ensure that the questions that were being posed in the user study were representative of the high-level tasks identified in Section 4.1. The feedback for the designs was positive and there was a large interest expressed in providing means of further exploration within the designs, mainly through the use of interaction and transitions. As the focus of this work was to validate the proposed diagram designs, improved exploration capability will be considered in future work.

5.2 Data Selection

The next step was to decide on a data set to use for the user study. For this, the data set from Greenberg *et al.* (2017) was used as all of the data used in their diagrams was available, the data set was currently the most similar data set to a FEW nexus data set used in a Sankey diagram, and the diagrams were manually created diagrams which presented what could be considered a best case representation of the data. Figure 5.1 shows a simplified example of two designs that represent data similar to that used in the user study.



a) **b)**

Figure 5.1: Example of two different designs presented in the user study: a) Traditional hybrid Sankey diagram vs b) NEST diagram design representation of a sample data set.

After deciding on this data set, a method was needed to select which of the states in the data set would be used for the user study. To do this, each of the states was first checked for node and edge consistency by subtracting the total amount of outflows from each of the nodes to make sure that they became zero. After doing this, there were twenty-six states left that appeared to have consistent data: Arizona, Connecticut, Delaware, Georgia, Hawaii, Idaho, Kentucky, Maine, Maryland, Massachusetts, Minnesota, Missouri, Nevada, New Hampshire, New Jersey, New York, North Carolina, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Vermont, Virginia, Washington, and Wisconsin.

In an attempt to select states from the remaining twenty-six states, pile sorting was first performed with several data visualization experts. However, it was quickly realized that this led to a large variation in what was considered the most important characteristics in a Sankey diagram. Although this could be an interesting study for future work, identifying a specific characteristic was not crucial to the task at hand. To avoid this variation, the states were instead grouped into clusters using k-means clustering. The input for the k-means clustering was the number of non-zero nodes and edges for each state. K-means clustering was done using between two and ten clusters to see which number of clusters appeared to best represent the data. The result of this can be seen in Figure 5.2.

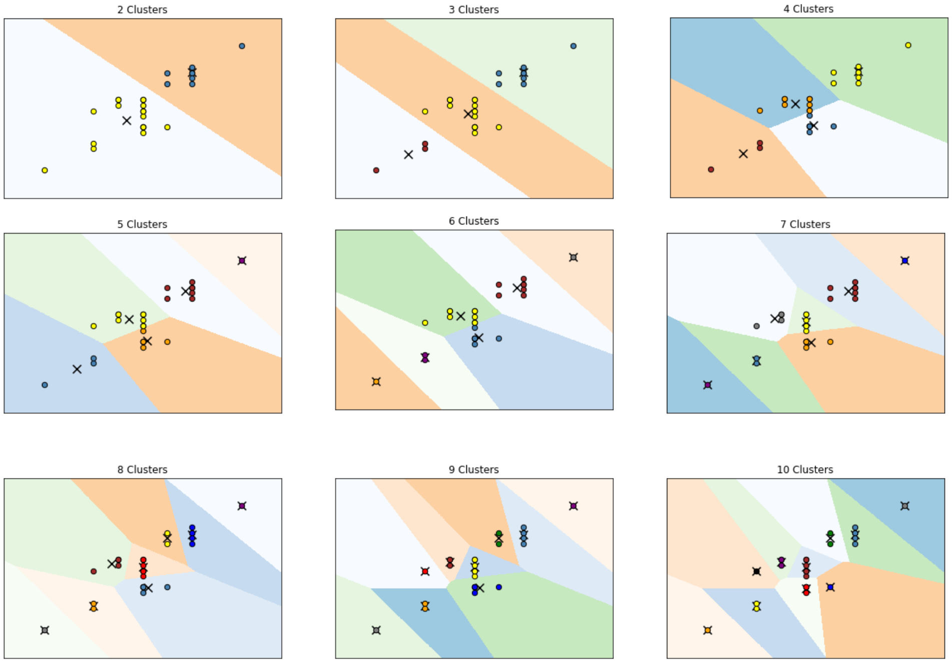


Figure 5.2: K-means clustering using between two and ten clusters. Data points are represented with a circle and cluster centers are represented with an x.

After performing k-means clustering, to narrow down the number of clusters that were considered from the k-means output, density-based spatial clustering of applications with noise (DBSCAN) was performed on the data set. This was done to try to minimize bias when choosing the number of clusters. Since the data set only contained 26 states, the epsilon was set to 0.2 and the minimum number of points to be considered a cluster was set to 2. The result of the DBSCAN can be seen in Figure 5.3.

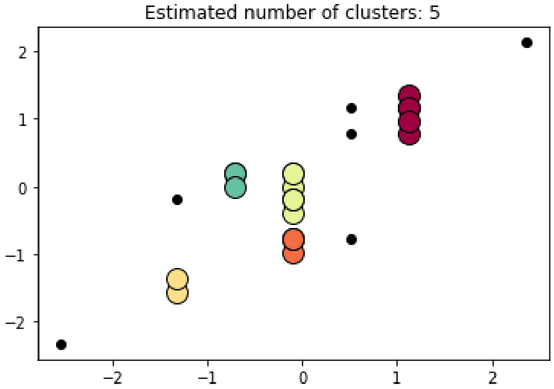


Figure 5.3: Clustering using DBSCAN with outliers represented in black.

After careful consideration, it was decided that seven clusters seemed to be the most accurate representation of the data in the k-means output. This cluster representation can be seen in Figure 5.4, compared to the five clusters that were identified using DBSCAN.

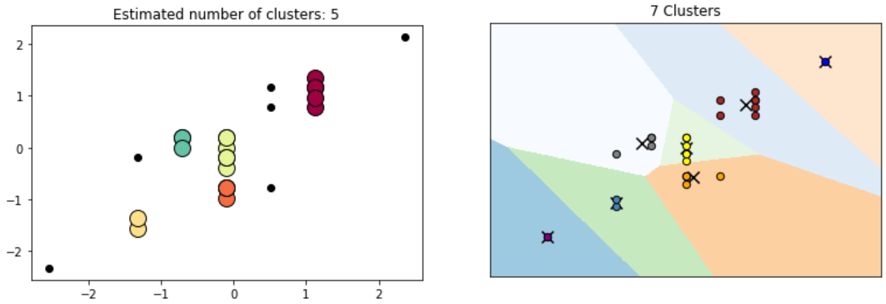


Figure 5.4: Clustering using DBSCAN on the left, with outliers represented in black, and k-means clustering using seven clusters on the right, with data points represented with a circle and cluster centers represented with an x.

It was then decided that a state from each of the clusters would be used in the user study, excluding the clusters that contained outliers found in the DBSCAN. The single states in the dark blue and light orange clusters, Vermont and Virginia, were discarded as they were considered outliers in the DBSCAN. This left what was considered five real clusters to be used. For the remaining five clusters, the states closest to the cluster centers were selected. Connecticut was selected from the dark orange cluster, Delaware was selected from the dark green cluster, Nevada was selected from the light green cluster, Tennessee was selected from the light blue cluster, and Wisconsin was selected from the white cluster.

5.3 Hypotheses

Examining the seven task taxonomy categories and twenty subcategories identified in Section 4.1, initial hypotheses were formed. These hypotheses fell into one of three categories: will perform better, will perform equivalent, or will perform worse than a traditional Sankey diagram, and were identified for each of the subcategories. The hypotheses for the subcategories, along with the reasoning, can be seen below.

5.3.1 Will Perform Better

T2.1 Amount of inflows/outflows of resources

- Increased ability to identify small inflows and outflows due to intermediate labels

T2.2 Amount of resource contributed from a specific parent node or supernode

- Increased ability to identify small inflows and outflows due to intermediate labels and uniform edges

T3.1 Identify parents of a single node or supernode

- Resources to the node/supernode are shown in the left column of a supernode
- Uniform edges allow small amounts to be traced

T3.2 Identify children of a single node or supernode

  - Resources to the node/supernode are shown in the right column of a supernode
  - Uniform edges allow small amounts to be traced

T3.4 Identify conversions from one resource into another resource

  - This is done in a separate row of the diagram

T4.1 Identify existence/absence of a path

  - Edges are uniform so small inflows and outflows can be traced

T4.2 Find a path of specific resource from the source to a potential final destination

  - Edges are uniform so small inflows and outflows can be traced

T4.3 Find all paths of specific resource from the final destination to the potential sources

  - Edges are uniform so small inflows and outflows can be traced
  - Artificial nodes allow destination flows to be routed to edge of diagram

T5.1 Intra-column understanding

  - Provides a qualitative and quantitative representation of the data
  - All inflows/outflows can be represented

T5.2 Patterns

  - Left and right columns of supernodes have numerical values, which can be used to compare multiple years/timesteps

T6.2 Identify maximum/minimum global resource type supply/usage

  - Qualitative and quantitative representations of the data are provided so multiple years/timesteps can be compared

T7.1 Inter-sector influence

  - Sectors are defined explicitly

5.3.2 Will Perform Equivalent

T1.1 Identify existence/absence of a node

  - Nodes are not significantly changed in the proposed diagrams

T1.2 Identify the name/type of a node

  - Nodes labeling and coloring are not significantly changed in the proposed diagrams

T1.3 Identify sectors used in a supernode

  - The original representation of sectors in a supernode is, for the most part, maintained in the proposed diagrams

T1.4 Count the number of supernodes with a given feature

  - Dependent on the feature, as the proposed diagrams have strengths and weaknesses for different features

T2.3 Amount of resource produced by a source node

  - Source nodes are not significantly changed in the proposed diagrams and occur in the same position as Sankey diagrams

T3.3 Identify splits/merges of different resource types

- Edge coloring is maintained in the proposed diagrams
- The proposed diagrams may be easier to identify small amounts of a resource being merged into other resource types

T6.1 Identify maximum/minimum local resource type supply/usage

- Maximum may become slower but minimum should be easier

5.3.3 Will Perform Worse

T7.2 Inter-sector resource usage

- Usage is combined to a single value (Water used for biomass production)
- Easier when finding total usages in a sector

5.4 Participants

To recruit participants for this user study, Amazon Mechanical Turk (MTurk) was utilized. Crowdsourcing was selected as the means of performing this user study as Heer and Bostock (2010) have examined its ability to be used to assess visualization designs, specifically using MTurk. Limitations presented in their work relevant to this user study, as well as limitations identified during the process of this user study, will be presented in Section 6.1. In total, 86 participants were recruited to participate in the user study. Each participants was paid $8 for their participation in the user study which lasted, in its entirety, approximately 50-60 minutes on average.

5.5 Procedure

For this experiment, there were a number of elements that went into the procedure. Since Sankey diagrams may not widely be known and new designs were being introduced, an introduction page explaining the characteristics of both Sankey diagrams and NEST diagram designs was first presented to each of the participants. Three introduction pages, one for each NEST diagram, were used to present the specific characteristics of each of the designs. Only the introduction page corresponding to the participants specific NEST diagram design was shown to the participant. To verify that the users had a sufficient understanding of the material presented in the introduction, a screening quiz was given. In this quiz, six questions were presented, two simple questions with feedback, and four more difficult questions without feedback. The first two questions were to allow the user to verify that they understood how the system worked, as well as provide the opportunity to answer simple questions related to Sankey and NEST diagrams. The remaining four questions were to verify that the participant had sufficient understanding of Sankey and NEST diagrams. Three of the four questions presented had to be answered correctly to participate in the real user study.

Next, the participant was presented with the real user study. At this stage, each participant was presented Sankey diagrams, as well as only one of the proposed NEST designs. The user study was conducted in this manner, as the Sankey diagram could be considered as a baseline and the primary focus was how the proposed diagrams compared to the Sankey diagrams, not to each other. Thirty unique questions were presented in total, once for each of the two diagram designs, with the ordering selected at random per question. The same question was presented once using one diagram followed by the same question for the other diagram. The ordering of the diagrams was randomly selected for each question. For each of the questions, states were pseudo-randomly selected from the five given states defined in Section 5.2 for each question, where the state selected for the Sankey diagram could not match the state selected for the NEST diagram for the same unique question. As different states were not considered as having a significant impact on the ability to answer the questions, this was only needed to minimize the impact of presenting the same question consecutively.

After answering the previous sixty questions, the participant was then presented with five questions asking about their preferred design and their perception of each of the diagrams in terms of accuracy and speed. These responses were not considered as a significant element during the evaluation of the results, as they only provide a persons preference, not an accurate means of quantifying their ability to use the diagram to solve problems. Finally, the participant was presented with an optional demographics survey to complete. Although the demographics information was not considered in the results of this user study, it did provide an idea of domain areas that may be interested in similar work.

5.6 Analysis Method

Results were analyzed for 86 participants: 30 for design 1, 23 for design 2, and 33 for design 3. The results were analyzed in terms of response time and accuracy. Since this user study was performed using MTurk, paired questions that had either the Sankey or NEST diagram's response time performance three standard deviations from the performance of the rest of the participants in that design study were discarded in terms of response time.

After removing outliers, response time and accuracy for each of the questions were examined using mean confidence intervals of 95%, and effect size using Cohen's d, which was introduced by Cohen (1988). In calculating the effect size using Cohens d, the difference between the means of the two samples is divided by the standard deviation of the data. When presenting the effect size, we considered the new effect size rules proposed by Sawilowsky (2009), which included the categories very small, very large, and huge, which were not categories in Cohens work. Although the new effect sizes were used for creating figures, to help separate values visually, the original effect sizes proposed by Cohen, with the large category being the greatest effect size, were considered during analysis. Cohen cautions using a strictly binning based approach and suggests that the significance of the effect size categories depends on the data. This lends well to the inclusion of confidence intervals when considering results. Using the categories originally proposed by Cohen seemed to best agree with what could be considered significant in the confidence intervals as well. The results that were considered to have a significant difference in performance were the questions with an effect size of large or above, which is a value of 0.50 or above, and those without a significant difference were the questions that had an effect size of medium or below, which is a value of below 0.50.

Further analysis was performed using a generalized linear mixed modeling (GLMM) Framework in SAS® 9.4's GLIMMIX procedure (Vonesh, 2012). This was used to determine the odds ratio between the Sankey diagram and each of the proposed NEST designs in both accuracy and response time. According to Bland and Altman (2000) an odds ratio is a way of representing probability, which is a form of representation most commonly used in betting. They further identify the following three reasons for their common use: 1) they can be used as an estimate for the relationship between binary variables 2) they allow for the examination of the effect of other variables using logistic regression 3) they have a convenient interpretation in case-control studies.

In this work, odds ratios were examined using a confidence interval of 95%. As the confidence interval for some of these ratios became quite large, and thus were hard to depict, the log-odds ratio is instead presented in the results. Using the log-odds ratio also created a symmetric depiction of the data across the x-axis rather than at $y = 1$ and allowed for the results that did not have an estimable odds ratio, or were not significantly different, to be presented as $y = 0$.

5.7 Results

Regarding task categories when examining the results for mean confidence intervals and effect size, the NEST design outperforms the Sankey diagrams in both accuracy and response time for value retrieval and link analysis questions. In extremum identification and inter-sector analysis, the NEST designs exhibit higher accuracy scores but slightly longer response time. In object identification and pattern recognition and column understanding, the differences are not significant in accuracy, while object identification questions exhibit shorter response time in the NEST designs and exhibit longer response time in pattern recognition and column understanding. The primary takeaway from these results is the significant improvement in accuracy using the NEST designs in value retrieval, link analysis, extremum identification, and inter-sector analysis questions, and improvement in response time in object identification questions.

The main weakness that is experienced in the NEST designs appears in questions 20, 21, and 22, which are path analysis questions. The main reason for the lower accuracy performance here is believed to be due to Sankey diagrams presenting a continuous flow from the beginning to the end, and thus less abstracted. However, in the NEST designs the links are only between neighboring supernodes, i.e. there are no internal links representing connections between nodes in a supernode, which could cause difficulty comprehending continuous flows. These results can be seen in Figures 5.5, 5.6, and 5.7.

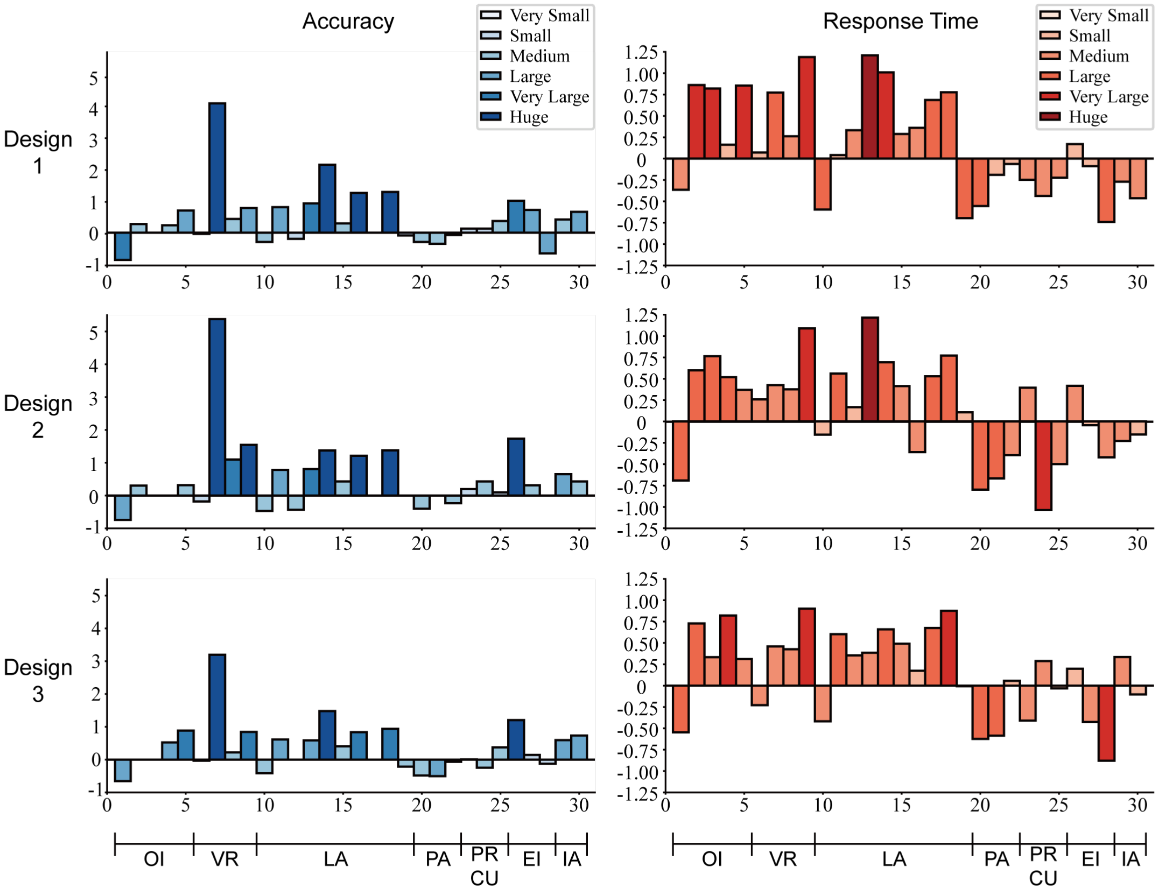


Figure 5.5: Effect size computed using Cohens d for the different NEST designs versus the Sankey diagram. Positive results indicate better performance in the NEST design and negative results indicate better performance in the Sankey diagram. The sign of the results for response time was flipped to allow for easier recognition of patterns between the accuracy and response time. OI, VR, LA, PA, PRCU, EI, and IA identify which questions were from which task category.

Another phenomenon identified from the results is that in the first question for each task category NEST designs showed lower performance in both accuracy and response time, including questions 1, 6, 10, and 23. This issue may indicate that the NEST designs suffer from a steeper learning curve, as this tends to disappear after the first few questions of the category. This is further supported by the lack of a low initial accuracy result for extremum identification tasks, which are similar to value retrieval tasks.

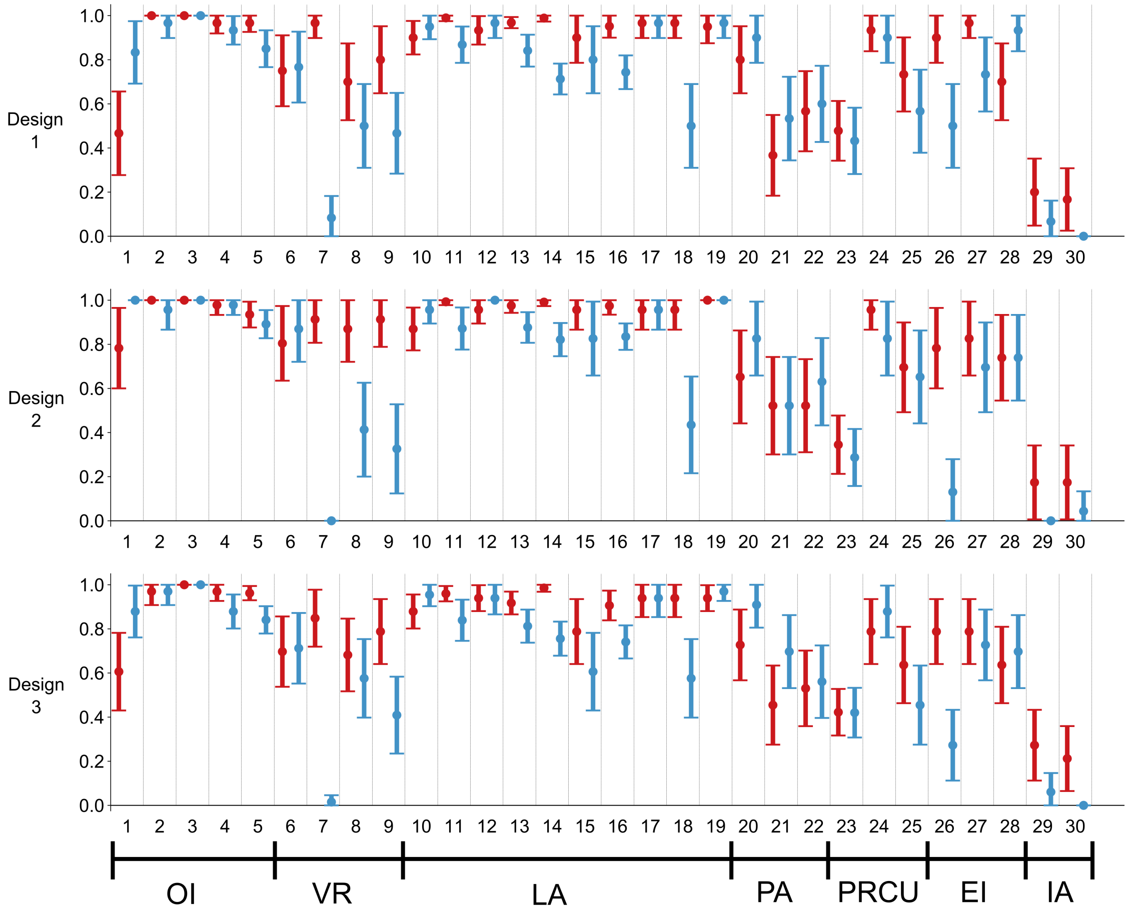


Figure 5.6: 95% Confidence intervals of mean accuracy values for each NEST design versus Sankey diagram. Red bars represent accuracy using NEST designs and blue ones represent corresponding Sankey diagrams. OI, VR, LA, PA, PRCU, EI, and IA identify which questions were from which task category.

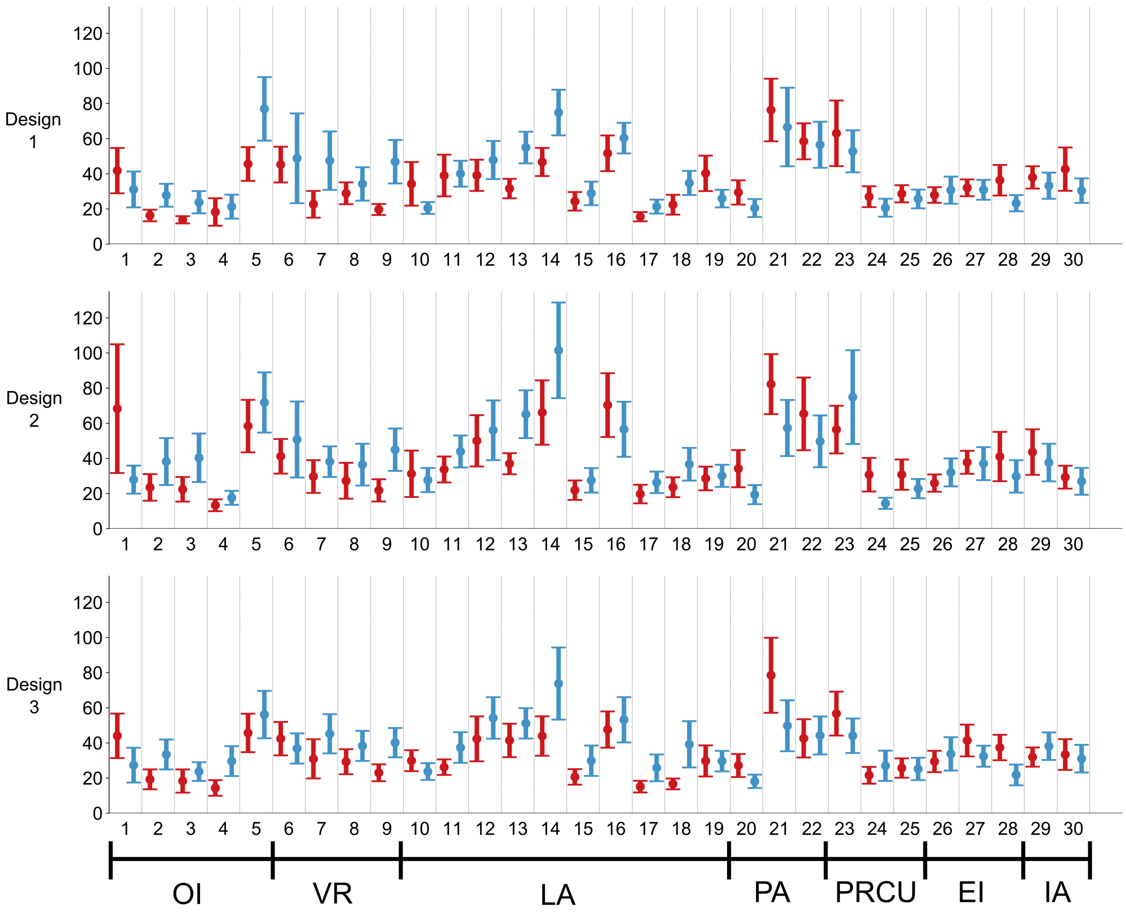


Figure 5.7: 95% Confidence intervals of mean response time for each NEST design versus Sankey diagram. Red bars represent response time using NEST designs and blue ones represent corresponding Sankey diagrams. OI, VR, LA, PA, PRCU, EI, and IA identify which questions were from which task category.

Similar trends can be seen in the analysis of the GLMM results in Figures 5.8 and 5.9. The primary difference is the reduction of the accuracy difference for the first question of each task category, which still occurs but is much less prevalent. The other primary difference is the better accuracy results for the NEST design in path analysis questions. Although the Sankey diagram does still perform better in this task category, the difference is much less noticeable, especially for NEST design 1. These differences are likely due to the answers being treated as binary, correct or incorrect, rather than discrete, like in the previous results.

Statistically, because of the structure of the study, a definitive conclusion cannot be drawn about the superiority of any one of the NEST designs to the other NEST design alternatives, as no participant was presented more than one of the NEST designs. Thus, the focus for the comparison of the NEST designs can solely be based on the odds ratios obtained, which could also be affected by other characteristics of the study.

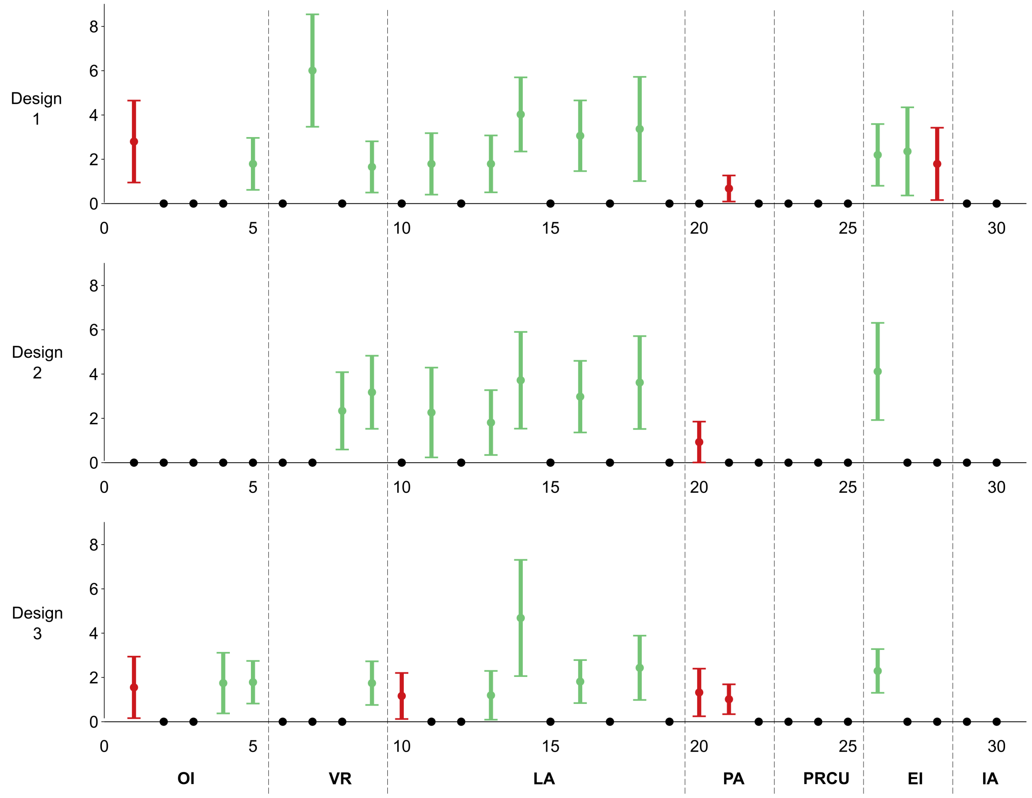


Figure 5.8: 95% confidence interval of the absolute value of log-odds ratio for accuracy. Red indicates better performance of Sankey diagram, green indicates better performance of NEST design, and black indicates no significant difference. OI, VR, LA, PA, PRCU, EI, and IA identify which questions were from which task category.

From these results, the odds ratio for accuracy in design 2 is higher than for design 1 and design 3, suggesting that participants tend to answer correctly for design 2 more often than for design 1 and design 3 when compared to the Sankey. Due to the limitations of the study, it is not possible however to state whether the difference between these odds ratios is significant. In terms of response time, the odds ratio for response time in design 3 is higher than for design 1 and design 2, suggesting that participants tend to answer more quickly for design 3 more often than for design 1 and design 2 when compared to the Sankey. Again, it cannot however be said that this is a significant difference. As ideally, accurate comprehension of the diagram should be more important than having a lower response time, from the odds ratio results, design 2 is recommended for visualizing heterogeneous data with directed flows.

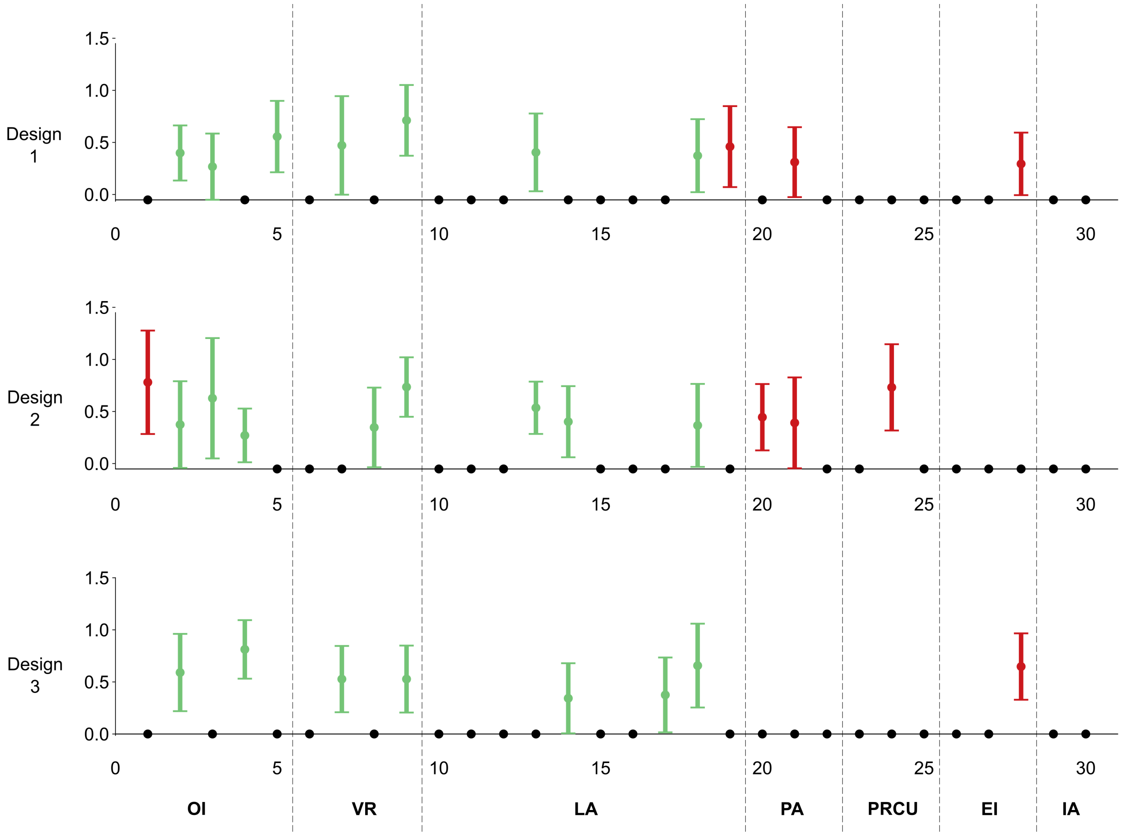


Figure 5.9: 95% confidence interval of the absolute value of log-odds ratio for response time. Red indicates better performance of Sankey diagram, green indicates better performance of NEST design, and black indicates no significant difference. OI, VR, LA, PA, PRCU, EI, and IA identify which questions were from which task category.

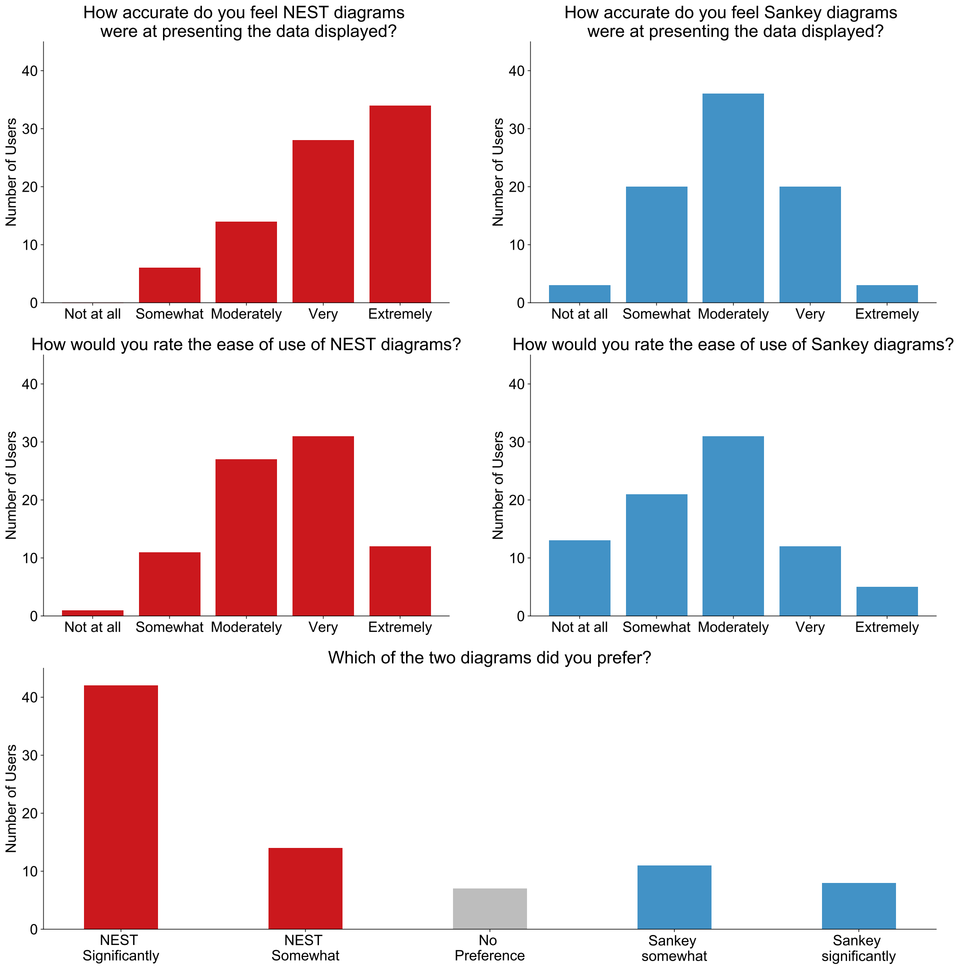


Figure 5.10: User preference results for the NEST and Sankey diagrams in terms of accuracy of presented data, ease of use, and preferred diagram.

Along with the user study, a user preference survey was also conducted. The results of this survey can be seen in Figure 5.10. The majority of the feedback received from the participants of the user study were in favor of the NEST designs over the Sankey diagram, mainly citing the ease of value retrieval tasks as their reasoning.

The negative feedback for the NEST designs were mainly focused toward difficulty in performing path-tracing tasks. This could potentially be from not providing enough information in the training to be able to fully understand path analysis tasks in the NEST designs, or potentially from the more abstracted flow of the resources in the NEST designs. Further research will be needed to evaluate which of these is occurring and how to minimize the issue in future work. The results from this survey did however provide areas of the design that needed improvement and identified areas that may need more focus in further interactive designs.

In the user preference survey, an optional section for additional comments was also provided to the participants. A reasonable number of participants, 34 of the 86, provided additional comments. The majority of these comments not only showed preference for the NEST designs but also many expressed interest in the user study itself. This highlights that crowdsourcing platforms can be an effective means of spreading awareness of and getting feedback on new visualization designs, due to its broad range of users. A sample of some of the comments received can be seen in the following:

- Diagram Feedback
  - "I liked the way that the NEST diagrams provided specific numbers for all components."
  - "The diagrams should use bolder paths to allow for easier use of tracking, but I did enjoy looking through the NEST diagrams."
  - "I found NEST easier to understand when there were quantities assigned to inputs and outputs. The Sankey was more clear when just trying to understand the flow of the inputs and outputs."

- “I strongly preferred NEST diagrams. It was very easy to see how resources were routed and consumed with a high level of detail. There were questions involving Sankey diagrams that left me confused or unsure because I could not see accurate numbers. They were easy to understand processes, though.”
- “I preferred the additional data the NEST diagrams provided over the “simplicity” of the Sankey diagrams.”
- “NEST was much easier to see discrete patterns. Sankey was easier to follow the flows from left to right, but became confused once there were small lines to follow.”
- “I did struggle more with the Sankey diagrams I felt on some questions, where it was hard to see such a fine line connecting items, but I also struggled to accurately read the NEST diagrams on other questions
- “I personally found the NEST diagrams to look like a bowl of spaghetti, but recognize they have more info. I would have preferred if there were numbers in the Sankey diagram to denote how much of one resource moved along, then it would have been perfect (although I guess they serve different purposes).”
- “The Sankey diagrams were confusing because some of the lines were hard to follow and some were really small so it was hard to see. The NEST diagrams were much easier because they had the numbers grouped into the boxes which took the guess work out as opposed to the Sankey. The lines did become a blur though when trying to follow them to an end point.”
- “Different questions could be answered more easily by different diagrams.”

  - “I felt the Nest diagram helped provide more information than the other diagram.”

- User Study Feedback
  - “I have never seen or heard of these diagrams before but it was fun to learn something new. Thanks for putting this together and making it available.”
  - “This was interesting. Appreciate you allowing me to take part!”
  - “I enjoyed this survey and reading both sets of diagrams.”
  - “Thank you! Interesting study.”
  - “This task is very interesting.”

Chapter 6

LIMITATIONS

6.1 User Study Limitations

As with nearly any user study, this user study came with a number of limitations. Limitations identified by Heer and Bostock (2010) when using MTurk that related to this user study were: insensitivity to factors like color blindness or limited visual acuity, need for qualification tasks, as well as clearly worded tasks with verifiable answers, and inability to collect fine-grained timing data. All of these factors could potentially have been experienced in this user study and several of the factors were very likely experienced.

In this experiment, participation in the user study was limited to individuals in the U.S. with a Human Intelligence Task (HIT) approval rate of 95% or above. This meant that of the tasks a participant had completed on MTurk, at least 95% of the tasks had to be accepted as valid participation, as defined by the task administer. Participants were also required to pass a preliminary screening with a score of 75% or above. This was to attempt to minimize the number of individuals that were "gaming" the system. Even with these restrictions, a number of results had to be discarded due to answering questions in an unreasonably small amount of time or providing duplicate answers that did not match the presented material. Having to enforce these restrictions filtered out a number of individuals that may potentially have had valid responses but were too similar to what was to be expected from a user "gaming" the system.

As a preliminary test was administered before an individual was able to participate in the user study, this enforced a constraint on the ability and knowledge-base of the participants. All users in this user study could be considered to be semi-experts in this area, or at least individuals with sufficient knowledge to understand the general concepts of Sankey diagrams and NEST diagrams. Limiting participation in such a way is a common issue that is faced in user study design, but is necessary to receive results that can capture the usability of the design rather than the ability of the participant.

Since this user study was run through MTurk, there were also no constraints placed on the computer hardware that was used during the study. This allowed for the possibility of issues of color blindness and limited visual acuity, which could potentially have hindered an individual's performance in this user study. Without constraining computer hardware, this could have potentially made the diagrams harder to see clearly. This however should have hindered performance using the NEST diagrams more than the Sankey diagrams as the elements in the Sankey diagram were larger than in the NEST diagrams and followed similar coloring schemes.

6.2 Sankey Diagram Limitations Identified from Previous Works

Sankey diagrams are generally used in the areas of energy and water systems and would appear to be a viable solution for FEW visualization as well. Using Sankey diagrams does, however, come with limitations, some of which exist in current Water-Energy Sankey diagrams. Bhaduri *et al.* (2018) state that Sankey diagrams are a labor-intensive visualization method and can often require large amounts of data, which may not be available, or may only be partially available. This can result in a lack of consistency within Sankey diagrams and can cause a misrepresentation of information since these diagrams are, for the most part, relatively simplistic. They illustrate three potential use cases for their proposed energy-water nexus knowledge discovery framework: energy-water-land interactions, state-level energy-water Sankey diagrams, and water demand and supply in shale oil and gas production. From this work, they further identify the data and analytic needs: access to observational and derived data sets, application of data and analytics at appropriate scales, facilitation of work outside the boundaries of individual models, and capabilities for the determination of future regional trajectories for climate, population, land use, economic activity, and energy technologies.

A large issue with using Sankey diagrams for complex systems is that they are often created manually. When produced manually these diagrams can better convey the intended information without a significant number of overlapping components, but the same cannot be said when automatically generating a Sankey diagram. Automatically generated Sankey diagrams have to be able to strategically layout components in such a way that visibility of all important components is maximized and interpretability by an end user is not compromised.

In a work by Alemasoom *et al.* (2016), they attempt to mitigate some of these issues by adding artificial, intermediate nodes. This was done to produce straighter edges, in an attempt to improve readability and reduce edge crossings. Automatically generated Sankey diagrams, for the most part, also do not handle components connected bidirectionally. In terms of the FEW nexus, this is a major limitation, as many components from different sectors interact with each other. This can be seen in terms of hydropower, where water is used for power generation and energy is used in water conveyance. Although a seemingly simple issue, the omittance of this information could present a very inaccurate view of the FEW nexus to an end user.

In a literature review by Soundararajan *et al.* (2014), just the energy sector was considered. In this review, they look at the energy transformation stage, which is the stage where the primary energy resources are converted to either energy carriers or secondary forms of energy. In some instances, the transformation and the supply stage are instead represented as a single entity, which causes a blurring of the different stages of an Energy Sankey diagram (Soundararajan *et al.*, 2014). This can result in an inaccurate understanding of how the flow between different components is occurring, and confusion in what resource is actually flowing.

Also, within the literature review by Soundararajan *et al.* (2014), they examine energy loss representation. They state that often energy losses that occur within the transportation and distribution stages of the energy flow are not represented. This could be an issue when looking at smaller scales, as improving the efficiency in these flows may have a significant impact on energy sustainability. The authors also go on to examine why energy loss is difficult to represent. The major issue being, what is considered as useful and what is considered as lost? They state that, for some applications, defining the useful energy and the lost energy is simple, but for other applications the definition becomes blurred. Although not nearly as difficult to define as it is within the energy sector, this issue does also occur within the water and food sectors as well.

The useful versus lost resource issue also carriers over to inter-sector resource use and can become quite difficult to define as well. Looking at a basic example, in terms of water use for cooling in energy, is water useful or is it lost? The answer to this becomes highly dependent on how the nexus is being considered. Looking only at the water sector, water was lost and there was no direct benefit that occurred, so this could be considered as loss within the water sector. Examining the water and energy sectors together, the water was used for energy generation, which can allow for the potential of increased conveyance and treatment of water, which could be considered as useful within the water sector. Although a basic example, this highlights issues that can occur within nexus research and decision support. This can also make FEW Sankey diagrams difficult, as these losses have to be conveyed to end users in such a way that it makes sense to a large number of individuals with diverse backgrounds.

When attempting to visualize the FEW nexus, the absolute magnitude of the production within the individual sectors can become an issue. When examining these sectors at a global scale in terms of mass, water withdrawal is roughly three orders of magnitude larger than consumption within the other sectors (Bijl *et al.*, 2018). This leads to the problem of presenting the connection between the individual sectors in a way such that none of the sectors are lost due to the characteristics of the other sectors. This problem is further compounded by the difference in units across the individual sectors. As there are different units in each sector, often a scale factor is used to present more similar values which can make the numerical understanding from the visualization more difficult. This also makes comparisons across similar visualizations more difficult, as this scale needs to be determined from the data that is given and may not necessarily lend itself well to different simulated scenarios where the water, energy, and food balance is no longer the same. Normalization can be used to minimize this issue but also introduces issues of its own, most significantly that values are then presented at a scale respective only to the diagram at hand.

6.3 Flow diagram Limitations Identified from Results

From the results of the user study, several limitations of flow diagrams, more specifically Sankey diagrams and NEST designs, for use in the visualization of the FEW nexus are identified. The most prominent of these issues is the performance experienced using both the NEST design and the Sankey diagram to answer inter-sector analysis questions. For these questions, using both the Sankey diagram and all NEST design alternatives, the mean confidence interval results for accuracy are around or below 20%. This highlights the participants inability to understand the complex interactions of the nexus represented in these designs. As the diagrams used in the user study only contained water and energy, this accuracy would likely have decreased fairly significantly with just the inclusion of the food sector, and would only continue to decrease with the inclusion of more sectors.

Similarly, low performance can be seen in value retrieval, path analysis, and pattern recognition and column understanding tasks, although not to the same degree as inter-sector analysis. Accuracy in these task categories were around 70%, 60%, and 60% respectively, which, although much higher than inter-sector analysis, is a much lower than desired result. This could likely highlight that both diagrams are still generally difficult to understand and would be more beneficial to experts than to non-experts. Several domain experts however also experienced difficulty using these diagrams to answer questions.

The final primary limitation is within Sankey diagram data sets. These data sets abstract a large amount of the detail of resource flows that could be important for detailed analysis. When looking at these data sets, after a resource passes through a supernode, you are no longer able to identify where the specific resource is going since it is now mixed with all other resources from the same sector. Whether this abstraction is due to not physically being able to identify this information or to attempting to simplify the information being presented to the end user, this could significantly hinder exploration capability that would be necessary for presenting a comprehensive view of the FEW nexus.

Chapter 7

CONCLUSIONS & FUTURE WORK

Providing a means of visualizing multi-sectoral resource data is imperative for successful collaboration when examining the FEW nexus. In this thesis, a design study of the Sankey diagram, one of the most common visualization approaches for critical infrastructure systems, was performed. This was accomplished by defining a set of visualization design requirements relating to Sankey diagrams, especially in terms of the FEW nexus, and developing a visualization approach, the NEST diagram, for representing multi-sectoral resource flows. The presented design requirements highlight the lack of research in evaluation methods in this area and provide a basis for the evaluation of flow diagrams in the future, which will be imperative for further nexus research. Although past research in design requirement specification has been done by Amar *et al.* (2005), Lee *et al.* (2006), Ahn *et al.* (2014), and Laha *et al.* (2015), these works focus on specification of generalized design requirements that do not fully capture the complexity of heterogeneous data with directed flow. Without preserving the spatial understanding of the underlying structure of the data being presented, making sense of resource flows can be difficult for non-experts that are often the end users of these kind of diagrams. This work expands on these works to provide design requirements for heterogeneous data with directed flows, targeted at the FEW nexus, but applicable to other resource flow data as well.

Further, a diagram design that has similar complexity to that of a Sankey diagram is presented to handle the highly complex interactions that occur within the FEW nexus. Improvements to design aesthetics have been proposed by Batini *et al.* (1986), Cui *et al.* (2008), Selassie *et al.* (2011), and Alemasoom *et al.* (2016), but these works focus on improving the impact of edges on the understandability of the diagram, rather that examining more elaborate design elements that can capture the information that may otherwise be lost. Rather than focusing solely on how edges cross, the NEST designs seek to offer alternative design improvements that maintain the simplicity of the Sankey diagram, which is beneficial to non-expert understanding of the data being presented, and also provide more detailed information that may be necessary for expert analysis of the underlying data.

In the user study, the ability to effectively use these diagram designs, compared to Sankey diagrams, has been demonstrated. This was accomplished by presenting diagrams with real world data relating to the Energy-Water nexus and asking questions that may be faced when examining this data. From the results, it was apparent that the NEST designs did not significantly decrease the accuracy or response rate of the participants for any of the task taxonomy categories, but did offer a significant increase in terms of accuracy and response rate for a number of the task taxonomy categories. These improvements can mainly be seen in value retrieval and link analysis questions, which can be regarded as key areas when examining heterogeneous data with directed flow. The NEST designs were further validated, in terms of preference, by comments received from both domain experts and the participants of the user study, many of which expressed a large interest in the NEST designs themselves, as well as interest in further exploration capability.

Future work for this thesis can be divided into two main areas: 1) overcoming new design challenges and 2) NEST design exploration capability. In this work, two primary challenges were identified when presenting the new NEST diagram designs: a steep initial learning curve and difficulty performing path analysis tasks. These challenges can be seen in the results for the first question of each task taxonomy category and the path analysis results. In these questions, an inconsistent dip in accuracy and a spike in response time occur. This is believed to be due to the NEST designs having a lot more data presented numerically and a more abstracted path of resources, from the source to the final destination. Further work in this area could not only help to improve the accuracy of results in design studies, but also help to identify effective strategies to introduce visualization designs to individuals that have not seen the design before. Future work in the exploration capability of the NEST designs includes identifying important elements of the NEST diagram designs, and flow diagrams in general, as well as integrating the NEST diagram into a visual analytics system. Identifying elements that can be considered important for data exploration in the NEST diagram would allow for more specific interaction capability that can highlight information that corresponds to the information in question. This research would also help to reduce adding exploration features that merely clutter the rest of the surrounding information without presenting anything of value to the end user. This exploration capability can then further be added to a visual analytics system to present a more comprehensive view of the FEW nexus, as well as provide a larger amount of detail to experts, that would otherwise not be possible using such a simplistic design.

REFERENCES

Ahn, J.-W., C. Plaisant and B. Shneiderman, “A task taxonomy for network evolution analysis”, IEEE Transactions on Visualization and Computer Graphics **20**, 3, 365–376 (2014).

Alemasoom, H., F. Samavati, J. Brosz and D. Layzell, “Energyviz: an interactive system for visualization of energy systems”, The Visual Computer **32**, 3, 403–413 (2016).

Amar, R., J. Eagan and J. Stasko, “Low-level components of analytic activity in information visualization”, in “IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.”, pp. 111–117 (2005).

Bartos, M. D. and M. V. Chester, “The conservation nexus: valuing interdependent water and energy savings in arizona”, Environmental science & technology **48**, 4 (2014).

Batini, C., E. Nardelli and R. Tamassia, “A layout algorithm for data flow diagrams”, IEEE Transactions on Software Engineering **SE-12**, 4, 538–546 (1986).

Berardy, A. and M. V. Chester, “Climate change vulnerability in the food, energy, and water nexus: concerns for agricultural production in arizona and its urban export supply”, Environmental Research Letters **12**, 3, 035004 (2017).

Bhaduri, B. L., A. Simon, M. R. Allen, J. Sanyal, R. N. Stewart and R. A. McManamay, “Energy-water nexus knowledge discovery framework, experts meeting report”, Tech. rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States) (2018).

Bijl, D. L., P. W. Bogaart, S. C. Dekker and D. P. van Vuuren, “Unpacking the nexus: Different spatial scales for water, food and energy”, Global Environmental Change **48**, 22–31 (2018).

Bland, J. M. and D. G. Altman, “The odds ratio”, BMJ **320**, 7247, 1468 (2000).

Bostock, M., V. Ogievetsky and J. Heer, “D3 data-driven documents”, IEEE Transactions on Visualization and Computer Graphics **17**, 12, 2301–2309 (2011).

Breitkreutz, B.-J., C. Stark and M. Tyers, “Osprey: a network visualization system”, Genome Biology **4**, 3, R22 (2003).

Cohen, J., “Statistical power analysis for the behavioural sciences hillsdale”, NJ: Lawrence Earlbaum Associates **2** (1988).

Cui, W., H. Zhou, H. Qu, P. C. Wong and X. Li, “Geometry-based edge clustering for graph visualization”, IEEE Transactions on Visualization and Computer Graphics **14**, 6, 1277–1284 (2008).

Curmi, E., R. Fenner, K. Richards, J. M. Allwood, B. Bajželj and G. M. Kopec, "Visualising a stochastic model of californian water resources using sankey diagrams", Water resources management **27**, 8, 3035–3050 (2013).

Dimara, E., A. Bezerianos and P. Dragicevic, "Conceptual and methodological issues in evaluating multidimensional visualizations for decision support", IEEE Transactions on Visualization and Computer Graphics **24**, 1, 749–759 (2018).

Eichelberger, H., "Nice class diagrams admit good design?", in "Proceedings of the 2003 ACM Symposium on Software Visualization", SoftVis '03, pp. 159–ff (ACM, New York, NY, USA, 2003).

Figueiras, A., "How to tell stories using visualization", in "2014 18th International Conference on Information Visualisation", pp. 18–18 (2014).

Gansner, E. R., Y. Hu, S. North and C. Scheidegger, "Multilevel agglomerative edge bundling for visualizing large graphs", in "2011 IEEE Pacific Visualization Symposium", pp. 187–194 (2011).

Gober, P. and C. W. Kirkwood, "Vulnerability assessment of climate-induced water shortage in phoenix", Proceedings of the National Academy of Sciences **107**, 50, 21295–21299 (2010).

Gotz, D. and M. X. Zhou, "Characterizing users' visual analytic activity for insight provenance", in "2008 IEEE Symposium on Visual Analytics Science and Technology", pp. 123–130 (2008).

Green, T. M., W. Ribarsky and B. Fisher, "Visual analytics for complex concepts using a human cognition model", in "2008 IEEE Symposium on Visual Analytics Science and Technology", pp. 91–98 (2008).

Greenberg, H. R., E. P. Shuster, A. Simon and S. L. Singer, *Development of Energy-Water Nexus State-level Hybrid Sankey Diagrams for 2010* (2017).

Hadhrawi, M. K., M. Nouh, A. Alfaris and A. Sanchez, "Copi: A web-based collaborative planning interface platform", in "Human Interface and the Management of Information. Information and Interaction for Learning, Culture, Collaboration and Business,", edited by S. Yamamoto, pp. 287–296 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013).

Harrison, L., K. Reinecke and R. Chang, "Infographic aesthetics: Designing for the first impression", in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", CHI '15, pp. 1187–1190 (ACM, New York, NY, USA, 2015).

Heer, J. and M. Bostock, "Crowdsourcing graphical perception: using mechanical turk to assess visualization design", in "Proceedings of the SIGCHI conference on human factors in computing systems", pp. 203–212 (ACM, 2010).

Holten, D. and J. J. Van Wijk, "Force-directed edge bundling for graph visualization", Computer Graphics Forum **28**, 3, 983–990 (2009).

Hoolohan, C., A. Larkin, C. McLachlan, R. Falconer, I. Soutar, J. Suckling, L. Varga, I. Haltas, A. Druckman, D. Lumbroso, M. Scott, D. Gilmour, R. Ledbetter, S. McGrane, C. Mitchell and D. Yu, "Engaging stakeholders in research to address water–energy–food (wef) nexus challenges", Sustainability Science (2018).

Inbar, O., N. Tractinsky and J. Meyer, "Minimalism in information visualization: Attitudes towards maximizing the data-ink ratio", in "Proceedings of the 14th European Conference on Cognitive Ergonomics: Invent! Explore!", ECCE '07, pp. 185–188 (ACM, New York, NY, USA, 2007).

Keim, D., G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer and G. Melançon, *Visual Analytics: Definition, Process, and Challenges*, pp. 154–175 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008).

Kerracher, N., J. Kennedy and K. Chalmers, "A task taxonomy for temporal graph visualisation", IEEE Transactions on Visualization and Computer Graphics **21**, 10, 1160–1172 (2015).

Laha, B., D. A. Bowman, D. H. Laidlaw and J. J. Socha, "A classification of user tasks in visual analysis of volume data", in "2015 IEEE Scientific Visualization Conference (SciVis)", pp. 1–8 (2015).

Lee, B., C. Plaisant, C. S. Parr, J.-D. Fekete and N. Henry, "Task taxonomy for graph visualization", in "Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization", BELIV '06, pp. 1–5 (ACM, New York, NY, USA, 2006).

Lukasczyk, J., X. Liang, W. Luo, E. D. Ragan, A. Middel, N. Bliss, D. White, H. Hagen and R. Maciejewski, "A Collaborative Web-Based Environmental Data Visualization and Analysis Framework", in "Workshop on Visualisation in Environmental Sciences (EnvirVis)", edited by A. Middel, K. Rink and G. H. Weber (The Eurographics Association, 2015).

Luo, S., C. Liu, B. Chen and K. Ma, "Ambiguity-free edge-bundling for interactive graph visualization", IEEE Transactions on Visualization and Computer Graphics **18**, 5, 810–821 (2012).

MacDonald, G. M., "Water, climate change, and sustainability in the southwest", Proceedings of the National Academy of Sciences **107**, 50, 21256–21262 (2010).

Martin, S., W. M. Brown, R. Klavans and K. W. Boyack, "Openord: an open-source toolbox for large graph layout", in "Visualization and Data Analysis", vol. 7868 (2011).

Mattor, K., M. Betsill, C. Huayhuaca, H. Huber-Stearns, T. Jedd, F. Sternlieb, P. Bixler, M. Luizza and A. S. Cheng, "Transdisciplinary research on environmental governance: A view from the inside", Environmental Science & Policy **42**, 90–100 (2014).

Miksch, S. and W. Aigner, “A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data”, Computers & Graphics **38**, 286–290 (2014).

Minard, C.-J., *Des Tableaux graphiques et des cartes figuratives, par M. Minard...* (Thunot, 1862).

Olmstead, S. M., “Climate change adaptation and water resource management: A review of the literature”, Energy Economics **46**, 500 – 509 (2014).

Purchase, H. C., J.-A. Allder and D. Carrington, “User preference of graph layout aesthetics: A uml study”, in “Graph Drawing”, edited by J. Marks, pp. 5–18 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2001).

Quintero, A., D. Konar and S. Pierre, “Prototyping an intelligent decision support system for improving urban infrastructures management”, European Journal of Operational Research **162**, 3, 654 – 672, decision-Aid to Improve Organisational Performance (2005).

Rinaldi, S. M., J. P. Peerenboom and T. K. Kelly, “Identifying, understanding, and analyzing critical infrastructure interdependencies”, IEEE Control Systems **21**, 6, 11–25 (2001).

Sawilowsky, S. S., “New effect size rules of thumb”, Journal of Modern Applied Statistical Methods pp. 597–599 (2009).

Segel, E. and J. Heer, “Narrative visualization: Telling stories with data”, IEEE Transactions on Visualization and Computer Graphics **16**, 6, 1139–1148 (2010).

Selassie, D., B. Heller and J. Heer, “Divided edge bundling for directional network data”, IEEE Transactions on Visualization and Computer Graphics **17**, 12, 2354–2363 (2011).

Shneiderman, B., “The eyes have it: a task by data type taxonomy for information visualizations”, in “Proceedings 1996 IEEE Symposium on Visual Languages”, pp. 336–343 (1996).

Soundararajan, K., H. K. Ho and B. Su, “Sankey diagram framework for energy and exergy flows”, Applied Energy **136**, 1035–1042 (2014).

Stasko, J. and E. Zhang, “Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations”, in “IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings”, pp. 57–65 (2000).

Sullivan, A., D. White, K. Larson and A. Wutich, “Towards water sensitive cities in the colorado river basin: A comparative historical analysis to inform future urban water sustainability transitions”, Sustainability **9**, 5 (2017).

Tanahashi, Y. and K. Ma, “Design considerations for optimizing storyline visualizations”, IEEE Transactions on Visualization and Computer Graphics **18**, 12, 2679–2688 (2012).

Tesarik, J., L. Dolezal and C. Kollmann, “User interface design practices in simple single page web applications”, in “2008 First International Conference on the Applications of Digital Information and Web Technologies (ICADIWT)”, pp. 223–228 (2008).

Tolone, W. J., “Interactive visualizations for critical infrastructure analysis”, International Journal of Critical Infrastructure Protection **2**, 3, 124 – 134 (2009).

Vonesh, E. F., *Generalized linear and nonlinear models for correlated data: theory and applications using SAS* (SAS Institute, 2012).

Wang, X., W. Dou, S.-E. Chen, W. Ribarsky and R. Chang, “An interactive visual analytics system for bridge management”, Computer Graphics Forum **29**, 3, 1033–1042 (2010).

Yang, Y. E. and S. Wi, “Informing regional water-energy-food nexus with system analysis and interactive visualization a case study in the great ruaha river of tanzania”, Agricultural Water Management **196**, 75–86 (2018).

Zhou, H., P. Xu, X. Yuan and H. Qu, “Edge bundling in information visualization”, Tsinghua Science and Technology **18**, 2, 145–156 (2013).

APPENDIX A

USER STUDY QUESTIONS

This section presents the questions and possible answers that were presented to the participants of the user study, separated into their respective task taxonomy categories.

Object Identification

1) Is there ocean discharge from the water sector?
   a. Yes
   b. No

2) What resource sector does the Geothermal node belong to, the energy sector or the water sector?
   a. Energy
   b. Water

3) What resource sector/sectors are used in the Transportation supernode? (Select all that apply)
   ☐ Energy
   ☐ Water

4) What resource sector/sectors are used in the Electricity Generation supernode? (Select all that apply)
   ☐ Energy
   ☐ Water

5) Which supernodes contain Fresh Surface water? (Select all that apply)
   ☐ Electricity Generation
   ☐ Public Water Supply
   ☐ Transportation
   ☐ Residential
   ☐ Commercial
   ☐ Industrial
   ☐ Irrigation
   ☐ Wastewater Treatment

Value Retrieval

6) What is the total inflow of water to the Irrigation supernode? (Enter -1 if do not have an estimate)

7) What is the inflow of Fresh Ground water to the Residential supernode? (Enter -1 if do not have an estimate)

8) What is the amount of energy contributed by the Petroleum node? (Enter -1 if do not have an estimate)

9) How much rejected energy comes from the Electricity Generation supernode? (Enter -1 if do not have an estimate)

☐

Link Analysis

10) Which supernodes contribute to the Commercial supernode? (Select all that apply)
- ☐ Electricity Generation
- ☐ Public Water Supply

11) Which supernodes use electricity? (Select all that apply)
- ☐ Public Water Supply
- ☐ Transportation
- ☐ Residential
- ☐ Commercial
- ☐ Industrial
- ☐ Irrigation
- ☐ Wastewater Treatment

12) After being processed (or consumed) in the Commercial supernode, what forms are the energy resources transformed into? (Select all that apply)
- ☐ Energy Services
- ☐ Rejected Energy
- ☐ Surface Discharge
- ☐ Ocean Discharge
- ☐ Consumed/Evaporated
- ☐ Injection
- ☐ Wastewater Treatment

13) Which resources are used to generate electricity? (Select all that apply)
- ☐ Petroleum
- ☐ Biomass
- ☐ Natural Gas
- ☐ Coal
- ☐ Nuclear
- ☐ Geothermal
- ☐ Hydro
- ☐ Wind/Solar
- ☐ Fresh Surface
- ☐ Saline Surface
- ☐ Fresh Ground
- ☐ Saline Ground

14) Which resources contribute to the Industrial supernode? (Select all that apply)
   - ☐ Petroleum
   - ☐ Biomass
   - ☐ Natural Gas
   - ☐ Coal
   - ☐ Nuclear
   - ☐ Geothermal
   - ☐ Hydro
   - ☐ Wind/Solar
   - ☐ Fresh Surface
   - ☐ Saline Surface
   - ☐ Fresh Ground
   - ☐ Saline Ground
   - ☐ Electricity
   - ☐ Public Water Supply

15) Is more water contributed to the Industrial supernode from the Public Water Supply or the Fresh Surface node directly?
   a. Public Water Supply
   b. Fresh Surface node

16) Which supernodes contribute to the Surface Discharge supernode? (Select all that apply)
   - ☐ Electricity Generation
   - ☐ Public Water Supply
   - ☐ Transportation
   - ☐ Residential
   - ☐ Commercial
   - ☐ Industrial
   - ☐ Irrigation
   - ☐ Wastewater Treatment

17) Is more water consumed in the Residential supernode from the Fresh Ground node or from the Public Water Supply?
   a. Fresh Ground node
   b. Public Water Supply

18) Does Wastewater Treatment contribute more energy to the Energy Services node or the Rejected Energy node?
   a. Energy Services node
   b. Rejected Energy node

19) What water sector Destination nodes are contributed to by water used in energy sector primary resource production? (Select all that apply)
   - ☐ Surface Discharge
   - ☐ Ocean Discharge
   - ☐ Consumed/Evaporated
   - ☐ Injection

Path Analysis

20) Is it possible that any of the water in the Surface Discharge node is from the Fresh Surface node?
   a. Yes
   b. No

21) Identify a path, starting from the Biomass node, that ends at the Rejected Energy node. List the supernodes encountered along the path, in order, by dragging from the Supernodes list to the Path list (Elements are placed at the end of the list. Selections can be moved back to the Supernodes list by dragging from the Path list to the Supernodes list.)

Supernodes	**Path**
Electricity Generation	
Public Water Supply	
Transportation	
Residential	
Commercial	
Industrial	
Irrigation	
Wastewater Treatment	

22) Identify a path from the Surface Discharge node to a potential resource node that contributed to it. List the supernodes encountered along the path, in order, by dragging from the Supernodes list to the Path list (Elements are placed at the end of the list. Selections can be moved back to the Supernodes list by dragging from the Path list to the Supernodes list.)

Supernodes	**Path**
Electricity Generation	
Public Water Supply	
Transportation	
Residential	
Commercial	
Industrial	
Irrigation	
Wastewater Treatment	

Pattern Recognition and Column Understanding

23) Which energy resources are used in at most one end use supernode? (Select all that apply)
    - ☐ Petroleum
    - ☐ Biomass
    - ☐ Natural Gas
    - ☐ Coal
    - ☐ Nuclear
    - ☐ Geothermal
    - ☐ Hydro
    - ☐ Wind/Solar

24) Sort the water resources by decreasing amount of usage. Which resource is in the 2nd position?
    a. Fresh Surface
    b. Saline Surface
    c. Fresh Ground
    d. Saline Ground

25) Sort the end-use supernodes by decreasing amount of Electricity usage. Which supernode is in the 3rd position?
    a. Transportation
    b. Residential
    c. Commercial
    d. Industrial
    e. Irrigation

Extremum Identification

26) Which end-use supernode used the smallest, non-zero, amount of Fresh Ground water?
    a. Transportation
    b. Residential
    c. Commercial
    d. Industrial
    e. Irrigation

27) Which end-use supernode used the largest amount of Natural Gas, considering both diagram 1 and diagram 2?
    a. Transportation
    b. Residential
    c. Commercial
    d. Industrial
    e. Irrigation

28) Which water sector Destination node contains the largest amount of total water, considering both diagram 1 and diagram 2?
   a. Surface Discharge
   b. Ocean Discharge
   c. Consumed/Evaporated
   d. Injection

Inter-sector Analysis

29) What is the total amount of water that is used for energy resource production? (Enter -1 if do not have an estimate)

30) How much of the total amount of water used by the energy sector is consumed or evaporated? (Enter -1 if do not have an estimate)

APPENDIX B

MEAN CONFIDENCE INTERVAL DATA

This section presents the mean confidence interval results for accuracy and response time for the 30 questions in each of the 3 design studies using the Sankey diagram and the respective NEST design. Accuracy results are presented as a score between 0 and 1, and response time results are presented in seconds.

	Design 1 Comparison		**Design 2 Comparison**		**Design 3 Comparison**	
Question	**Mean Accuracy Confidence Interval for NEST Design 1**	**Mean Accuracy Confidence Interval for Sankey Diagram**	**Mean Accuracy Confidence Interval for NEST Design 2**	**Mean Accuracy Confidence Interval for Sankey Diagram**	**Mean Accuracy Confidence Interval for NEST Design 3**	**Mean Accuracy Confidence Interval for Sankey Diagram**
1	0.28- 0.66	0.69 - 0.97	0.60 - 0.97	1.0 - 1.0	0.43 - 0.78	0.76 - 1.0
2	1.0 - 1.0	0.90 - 1.00	1.0 - 1.0	0.87 - 1.0	0.91 - 1.0	0.91 - 1.0
3	1.0 - 1.0	1.0 - 1.0	1.0 - 1.0	1.0 - 1.0	1.0 - 1.0	1.0 - 1.0
4	0.92- 1.0	0.87 – 1.0	0.93 - 1.0	0.93 - 1.0	0.93 - 1.0	0.80 - 0.96
5	0.93 - 1.0	0.77 - 0.93	0.88 - 0.99	0.83 - 0.96	0.93 - 0.99	0.78 - 0.90
6	0.59 - 0.91	0.61 - 0.93	0.64 - 0.97	0.72 - 1.0	0.54 - 0.86	0.55 - 0.87
7	0.90 - 1.0	0 - 0.18	0.81 - 1.0	0.0 - 0.0	0.72 - 0.98	0 - 0.05
8	0.53 - 0.87	0.31 - 0.69	0.72 - 1.0	0.20 - 0.63	0.5171 - 0.8466	0.40 - 0.75
9	0.65 - 0.95	0.28 - 0.65	0.79 - 1.0	0.12 - 0.53	0.64 - 0.94	0.23 - 0.58
10	0.82 - 0.98	0.89 - 1.0	0.77 - 0.97	0.89 - 1.0	0.80 - 0.96	0.90 - 1.0
11	0.98 - 1.0	0.789 - 0.95	0.98 - 1.0	0.78 - 0.97	0.92 - 0.99	0.75 - 0.93
12	0.87 - 1.0	0.90 - 1.0	0.89 - 1.0	1.0 - 1.0	0.88 – 1.0	0.87 - 1.0
13	0.94 - 0.99	0.77 - 0.91	0.94 - 1.0	0.81 - 0.95	0.87 - 0.97	0.74 - 0.89
14	0.97 - 1.0	0.64 - 0.78	0.97 - 1.0	0.75 - 0.90	0.97 - 1.0	0.68 - 0.83
15	0.79 - 1.0	0.65 - 0.95	0.87 - 1.0	0.66 - 0.99	0.64 - 0.94	0.43 - 0.78
16	0.90 - 1.0	0.67 - 0.82	0.93 - 1.0	0.77 - 0.89	0.84 - 0.97	0.67 - 0.82
17	0.90 - 1.0	0.90 - 1.0	0.87 - 1.0	0.867 - 1.0	0.85 - 1.0	0.85 - 1.0
18	0.90 - 1.0	0.31 - 0.69	0.87 - 1.0	0.22 - 0.65	0.85 - 1.0	0.70 - 0.75
19	0.87 - 1.0	0.90 - 1.0	1.0 - 1.0	1.0 - 1.0	0.88 – 1.0	0.93 - 1.0
20	0.65 - 0.95	0.79 - 1.0	0.44 - 0.86	0.66 - 0.99	0.57 - 0.89	0.81 - 1.0
21	0.18 - 0.55	0.34 - 0.72	0.30 - 0.74	0.30 - 0.74	0.28 - 0.63	0.53 - 0.86
22	0.39 - 0.75	0.43 - 0.77	0.31 - 0.73	0.43 - 0.83	0.36 - 0.70	0.40 - 0.73
23	0.34 - 0.61	0.28 - 0.58	0.21 - 0.48	0.16 - 0.42	0.32 - 0.53	0.31 - 0.53
24	0.84 - 1.00	0.79 - 1.0	0.87 - 1.0	0.66 - 0.99	0.64 - 0.94	0.76 - 1.0
25	0.57 - 0.90	0.39 - 0.75	0.49 - 0.90	0.44 - 0.86	0.46 - 0.81	0.28 - 0.63
26	0.79 - 1.0	0.31 - 0.69	0.60 - 0.97	0 - 0.28	0.64 - 0.94	0.11 - 0.43
27	0.90 - 1.0	0.57 - 0.90	0.66 - 0.99	0.49 - 0.90	0.64 - 0.94	0.57 - 0.89
28	0.53 - 0.87	0.84 - 1.0	0.55 - 0.93	0.55 - 0.93	0.46 - 0.81	0.531 - 0.86
29	0.05 - 0.35	0 - 0.16	0.01 - 0.34	0.0 - 0.0	0.11 - 0.43	0 - 0.15
30	0.03 - 0.31	0.0 - 0.0	0.01 - 0.34	0 - 0.13	0.06 - 0.36	0.0 - 0.0

*Values are rounded to 2 decimal places

Question	Design 1 Comparison		Design 2 Comparison		Design 3 Comparison	
	Mean Response Time Confidence Interval for NEST Design 1	Mean Response Time Confidence Interval for Sankey Diagram	Mean Response Time Confidence Interval for NEST Design 2	Mean Response Time Confidence Interval for Sankey Diagram	Mean Response Time Confidence Interval for NEST Design 3	Mean Response Time Confidence Interval for Sankey Diagram
1	28.9 - 54.7	20.9 - 41.3	31.7 - 104.9	19.9 - 35.9	31.4 - 56.7	17.5 - 37.3
2	13.0 - 19.6	21.3 - 34.3	15.9 - 31.1	24.8 - 51.6	13.6 - 25.0	24.9 - 42.0
3	11.8 - 15.9	17.5 - 30.2	15.4 - 29.4	26.5 - 54.2	11.7 - 24.9	18.4 - 29.0
4	10.4 - 26.1	14.5 - 28.1	9.9 - 16.7	13.6 - 21.5	9.9 - 18.8	21.1 - 38.2
5	36.0 - 55.2	58.9 - 95.1	43.5 - 73.3	54.7 - 89.0	34.7 - 56.7	42.7 - 69.6
6	35.1 - 55.4	23.3 - 74.4	31.3 - 51.0	29.1 - 72.4	32.9 - 52.0	28.3 - 45.5
7	15.0 - 30.3	30.8 - 64.1	20.3 - 39.0	29.4 - 46.8	19.8 - 42.1	34.1 - 56.4
8	22.7 - 35.1	24.7 - 43.7	17.0 - 37.5	24.5 - 48.3	22.2 - 36.4	29.8 - 46.9
9	16.5 - 22.8	34.5 - 59.3	15.5 - 28.1	32.9 - 57.0	18.2 - 27.9	31.8 - 48.5
10	21.9 - 46.7	17.2 - 23.9	18.0 - 44.4	20.8 - 34.6	23.9 - 35.9	18.9 - 28.5
11	27.1 - 50.9	32.7 - 47.4	26.3 - 41.1	34.8 - 53.0	21.8 - 30.7	28.7 - 46.1
12	30.3 - 48.1	37.0 - 58.7	35.4 - 64.6	39.0 - 73.0	29.6 - 55.1	42.4 - 66.1
13	26.1 - 37.1	46.0 - 64.0	31.0 - 43.0	51.5 - 78.8	31.9 - 51.0	42.5 - 59.8
14	38.7 - 54.7	61.9 - 87.9	47.8 - 84.4	74.3 -128.7	32.7 - 55.2	53.3 - 94.3
15	19.1 - 29.6	22.1 - 35.6	16.4 - 27.4	20.5 - 34.5	16.2 - 25.1	21.2 - 38.5
16	41.5 - 61.8	51.6 - 69.0	52.2 - 88.5	40.9 - 72.2	37.3 - 58.0	40.3 - 66.1
17	13.0 - 18.2	17.3 - 25.3	14.4 - 25.0	20.2 - 32.5	11.8 - 18.5	18.2 - 33.4
18	16.8 - 28.1	27.9 - 41.7	17.9 - 29.2	27.4 - 46.0	13.6 - 19.8	26.0 - 52.4
19	30.2 - 50.4	20.9 - 30.9	21.8 - 35.3	23.8 - 36.4	20.9 - 38.6	23.8 - 35.5
20	22.5 - 36.3	15.4 - 25.6	23.6 - 44.7	13.9 - 24.8	20.6 - 33.7	14.3 - 22.0
21	58.5 - 94.1	44.3 - 89.0	65.2 - 99.3	41.3 - 73.3	57.2 - 99.9	35.3 - 64.3
22	48.2 - 68.8	43.4 - 69.7	44.7 - 86.0	35.0 - 64.5	31.7 - 53.5	33.4 - 55.1
23	44.4 - 81.7	40.8 - 64.8	42.8 - 70.0	48.2 - 101.6	44.2 - 69.2	34.3 - 53.9
24	21.0 - 32.9	15.6 - 25.8	21.1 - 40.3	11.2 - 17.6	16.8 - 26.4	18.4 - 35.6
25	23.7 - 33.5	20.4 - 31.0	22.1 - 39.4	17.3 - 28.3	20.2 - 31.2	18.8 - 31.6
26	23.5 - 32.4	23.0 - 38.4	21.0 - 30.8	24.0 - 40.0	23.4 - 35.5	24.3 - 43.3
27	27.2 - 36.8	25.2 - 36.5	31.2 - 44.3	27.6 - 46.4	32.3 - 50.5	26.5 - 38.5
28	27.6 - 45.1	18.7 - 27.9	27.0 - 55.1	20.5 – 39.0	30.2 - 44.6	15.9 - 27.7
29	31.6 - 44.3	25.8 - 40.7	30.6 - 56.6	26.9 - 48.3	26.5 - 37.4	30.3 - 46.0
30	30.3 - 55.0	23.4 - 37.4	22.8 - 35.8	19.3 - 34.6	24.6 - 42.2	23.2 - 38.9

*Values are rounded to 1 decimal place

APPENDIX C

EFFECT SIZE DATA

This section presents the effect size results, using Cohen's d, for accuracy and response time for the 30 questions in each of the 3 design studies using the Sankey diagram and the respective NEST design. Positive accuracy results present better performance using the NEST design and negative results present better performance using the Sankey diagram. Negative response time results present better performance using the NEST design and positive results present better performance using the Sankey diagram.

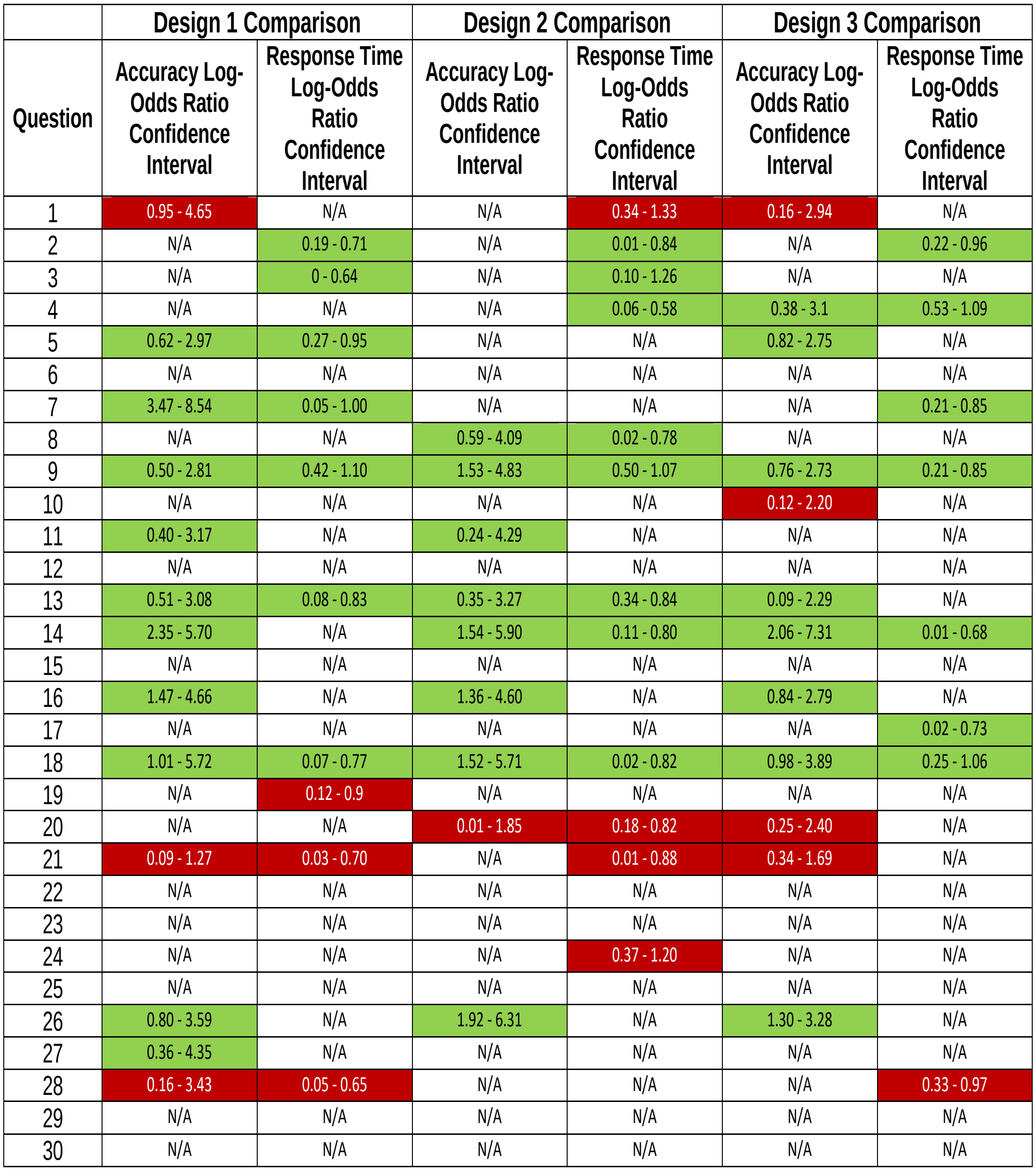
	Design 1 Comparison		**Design 2 Comparison**		**Design 3 Comparison**	
Question	**Accuracy Effect Size for NEST Design 1 vs Sankey Diagram**	**Response Time Effect Size for NEST Design 1 vs Sankey Diagram**	**Accuracy Effect Size for NEST Design 2 vs Sankey Diagram**	**Response Time Effect Size for NEST Design 2 vs Sankey Diagram**	**Accuracy Effect Size for NEST Design 3 vs Sankey Diagram**	**Response Time Effect Size for NEST Design 3 vs Sankey Diagram**
1	-0.83	0.36	-0.75	0.70	-0.66	0.55
2	0.26	-0.86	0.30	-0.60	0.0	-0.73
3	0.0	-0.82	0.0	-0.76	0.0	-0.33
4	0.22	-0.16	0.0	-0.52	0.52	-0.82
5	0.68	-0.86	0.31	-0.37	0.88	-0.31
6	-0.04	-0.07	-0.18	-0.26	-0.03	0.23
7	3.94	-0.77	5.38	-0.43	3.20	-0.46
8	0.42	-0.26	1.10	-0.38	0.22	-0.43
9	0.75	-1.19	1.55	-1.09	0.85	-0.90
10	-0.28	0.60	-0.47	0.15	-0.42	0.42
11	0.78	-0.04	0.78	-0.56	0.62	-0.60
12	-0.19	-0.33	-0.44	-0.17	0.0	-0.35
13	0.89	-1.21	0.81	-1.22	0.59	-0.39
14	2.07	-1.01	1.37	-0.69	1.48	-0.66
15	0.28	-0.29	0.43	-0.41	0.40	-0.49
16	1.21	-0.36	1.21	0.36	0.83	-0.17
17	0.0	-0.69	0.0	-0.53	0.0	-0.68
18	1.24	-0.78	1.38	-0.77	0.94	-0.88
19	-0.09	0.70	0.0	-0.11	-0.21	0.01
20	-0.28	0.56	-0.40	0.80	-0.49	0.62
21	-0.34	0.19	0.0	0.67	-0.51	0.59
22	-0.07	0.06	-0.23	0.40	-0.07	-0.06
23	0.12	0.25	0.20	-0.40	0.01	0.41
24	0.12	0.44	0.43	1.04	-0.25	-0.29
25	0.35	0.22	0.09	0.50	0.37	0.03
26	0.97	-0.17	1.73	-0.42	1.21	-0.20
27	0.69	0.09	0.31	0.04	0.14	0.43
28	-0.63	0.74	0.0	0.42	-0.13	0.88
29	0.4	0.27	0.65	0.23	0.59	-0.34
30	0.63	0.46	0.43	0.15	0.73	0.10

*Values are rounded to 2 decimal places

APPENDIX D

LOG-ODDS RATIO CONFIDENCE INTERVAL DATA

This section presents the log-odds ratio confidence interval results for accuracy and response time for the 30 questions in each of the 3 design studies using the Sankey diagram and the respective NEST design. Elements highlighted in red present significantly better performance using the Sankey diagram and elements highlighted in green present significantly better performance using the NEST design. N/A results present no significant difference in performance between the design alternatives and the Sankey diagram.



	Design 1 Comparison		Design 2 Comparison		Design 3 Comparison	
Question	Accuracy Log-Odds Ratio Confidence Interval	Response Time Log-Odds Ratio Confidence Interval	Accuracy Log-Odds Ratio Confidence Interval	Response Time Log-Odds Ratio Confidence Interval	Accuracy Log-Odds Ratio Confidence Interval	Response Time Log-Odds Ratio Confidence Interval
1	0.95 - 4.65	N/A	N/A	0.34 - 1.33	0.16 - 2.94	N/A
2	N/A	0.19 - 0.71	N/A	0.01 - 0.84	N/A	0.22 - 0.96
3	N/A	0 - 0.64	N/A	0.10 - 1.26	N/A	N/A
4	N/A	N/A	N/A	0.06 - 0.58	0.38 - 3.1	0.53 - 1.09
5	0.62 - 2.97	0.27 - 0.95	N/A	N/A	0.82 - 2.75	N/A
6	N/A	N/A	N/A	N/A	N/A	N/A
7	3.47 - 8.54	0.05 - 1.00	N/A	N/A	N/A	0.21 - 0.85
8	N/A	N/A	0.59 - 4.09	0.02 - 0.78	N/A	N/A
9	0.50 - 2.81	0.42 - 1.10	1.53 - 4.83	0.50 - 1.07	0.76 - 2.73	0.21 - 0.85
10	N/A	N/A	N/A	N/A	0.12 - 2.20	N/A
11	0.40 - 3.17	N/A	0.24 - 4.29	N/A	N/A	N/A
12	N/A	N/A	N/A	N/A	N/A	N/A
13	0.51 - 3.08	0.08 - 0.83	0.35 - 3.27	0.34 - 0.84	0.09 - 2.29	N/A
14	2.35 - 5.70	N/A	1.54 - 5.90	0.11 - 0.80	2.06 - 7.31	0.01 - 0.68
15	N/A	N/A	N/A	N/A	N/A	N/A
16	1.47 - 4.66	N/A	1.36 - 4.60	N/A	0.84 - 2.79	N/A
17	N/A	N/A	N/A	N/A	N/A	0.02 - 0.73
18	1.01 - 5.72	0.07 - 0.77	1.52 - 5.71	0.02 - 0.82	0.98 - 3.89	0.25 - 1.06
19	N/A	0.12 - 0.9	N/A	N/A	N/A	N/A
20	N/A	N/A	0.01 - 1.85	0.18 - 0.82	0.25 - 2.40	N/A
21	0.09 - 1.27	0.03 - 0.70	N/A	0.01 - 0.88	0.34 - 1.69	N/A
22	N/A	N/A	N/A	N/A	N/A	N/A
23	N/A	N/A	N/A	N/A	N/A	N/A
24	N/A	N/A	N/A	0.37 - 1.20	N/A	N/A
25	N/A	N/A	N/A	N/A	N/A	N/A
26	0.80 - 3.59	N/A	1.92 - 6.31	N/A	1.30 - 3.28	N/A
27	0.36 - 4.35	N/A	N/A	N/A	N/A	N/A
28	0.16 - 3.43	0.05 - 0.65	N/A	N/A	N/A	0.33 - 0.97
29	N/A	N/A	N/A	N/A	N/A	N/A
30	N/A	N/A	N/A	N/A	N/A	N/A

*Values are rounded to 2 decimal places