**Identifying stability, topological significance, and redundancies in water resource networks using parallel coordinate plotting**

Leah Meeks, S.M. ASCE, M.E., EIT1 and David E. Rosenberg, M. ASCE Ph.D2

Graduate Researcher, Dept. of Civil & Env. Engineering and Utah Water Research Laboratory, Utah State University Dept. of Civil & Env. Engineering, 4110 Old Main Hill, Utah State University, Logan, UT 84322-4110; c:(435)-881-8697, [l.meeks@aggiemail.usu.edu](mailto:l.meeks@aggiemail.usu.edu);

2 Assistant Professor Dept. of Civil & Env. Engineering and Utah Water Research Laboratory, Utah State University Dept. of Civil & Env. Engineering, 4110 Old Main Hill, Utah State University, Logan, UT 84322-4110; o:(435)-797-8689, f:(435)-797-1185, [david.rosenberg@usu.edu](mailto:david.rosenberg@usu.edu).

**Abstract**

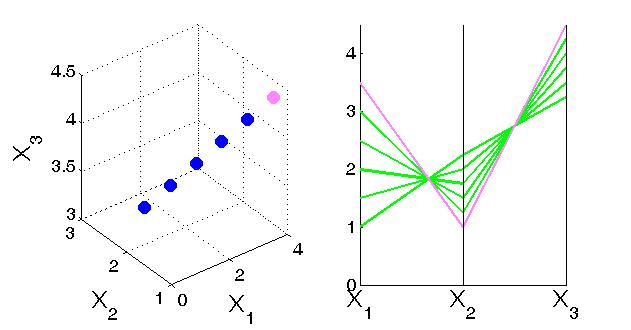
The size and complexity of water resources networks typically require a large number of computationally intensive simulations to test effects of changes in network structure or management. Current tools can only visualize the effects of a few changes. Here, we introduce a new method and tool that uses parallel coordinate plotting to simultaneously visualize large water resources networks plus identify and rank nodes that are (1) stable (their connectivity does not depends on the existence of particular nodes), (2) topologically significant (when removed or added to the network, they cause other nodes to be unstable), and (3) redundant (a node pair that has similar connections). The tool works by calculating node centrality, creating a parallel coordinate plot, calculating pairwise differences between the elements comprising each plot trace, and calculating each node’s stability, topological significance, and redundancy. We apply the tool to the 56-node lower Bear River water system that stretches from southern Idaho to the Great Salt Lake, Utah. Nodes that are connected to only one other node are the least stable, including Great Salt Lake, Malad River, and Evaporation from Hyrum Reservoir. The three most topologically significant nodes are Cutler and the two junctions connecting the South Cache Valley and the Weber branches to the rest of the network. There are five highly redundant node pairs with over 96% of the same connections including the Cache Valley Irrigation and Cache Valley New Municipal and Industrial service areas. These results suggest that the redundant Cache Valley Irrigation service area is a promising source to transfer water from agriculture to urban use. The New Box Elder County Irrigation and South Cache Irrigation service areas have very low topological significance ranks and suggest that these irrigation areas may also be promising sources of water transfers. Results also identify candidate locations to (i) remove dams (reservoirs with low topological significance or high redundancy), (ii) implement conservation measures, develop new alternative supplies, or monitor flows (unstable nodes), and (iii) protect environmental and ecological services (nodes with high topological significance). Future work should incorporate flow direction and magnitude. The tool scales to very large networks and identifies the most promising nodes to subsequently focus computationally-intensive simulation and sensitivity analysis efforts.

**CE Database subject headings:** Parallel coordinates, Water management, Bear River, Network analysis

**Introduction**

Water resources networks are often large and have numerous water supply, reservoir, diversion, and demand site nodes that are connected through natural and engineered conveyance links. Their large size and complexity typically require analysts to make numerous and computationally intensive simulation runs to test the effects of particular changes in network structure or management (Louchs et al., 2005). Current simulation-based network analysis tools can only visualize the effects of two or a few network changes (Wegman, 1990). Here, we develop a new method and tool that both manipulates the connectivity matrix representing the network topology and uses parallel coordinate plotting to automatically identify and rank the key nodes in a network. As a result, the analysis can suggest improvements to the water system configuration and management.

Parallel coordinate plots (PCP) can simultaneously display a large number of dimensions and are a way to visualize large networks (Singer and Greeenshpan, 2009). PCPs draw axes in parallel and allow the viewer to see a very large number of dimensions side-by-side in one figure within the two-dimensional confines of a printed page or screen (Inselberg, 1985). PCPs contrast with traditional Cartesian coordinate plots which plot axes orthogonally and can only illustrate two or three dimensions. A point in three-dimensional orthogonal coordinates is represented as a vector in parallel coordinates whose elements are connected by a trace of polygonal lines (Figure 1).



a

b

b

Figure 1. Comparison of data in (a) Cartesian coordinates and (b) parallel coordinates (Rosenberg, 2012).

PCPs can improve understanding of systems because of their capacity to display a large number of dimensions simultaneously. Albazzaz et al. (2005) used PCP to detect abnormal operations in wastewater treatment by tracking 38 variables for 527 days; they found 17 abnormal days. PCPs were used to identify errors in raw data through PCP visual inspection for lines that did not follow the trends (Edsall, 2003).

Choi et al (2008) applied PCPs to detect unknown large-scale attacks through internet sources. The plots allowed for visual determination of anomalies in source internet protocol (IP) address, destination IP address, destination port, and average flow packet length of internet network traffic to classify attacks.

PCPs have been dynamically linked to geographical maps to help with analysis of geographic phenomena (Edsall 2003). Geographic maps or scatter plots are dynamically updated with values from the PCP (Andrienko and Andrienko, 2001). Edsall (2003) used PCPs to expand the variables used to classify hurricanes beyond sustained wind speed to include water vapor content, sea surface temperature, and warning systems. In addition, PCPs helped identify variables, such as barometric pressure, after a hurricane event that had extreme values. PCPs also helped compare multiple variables at a particular latitude.

Singer and Greenshpan (2009) propose a method using PCPs to analyze networks called node extraction visualization (NEVIS) that is included in Inselberg’s compilation textbook on parallel coordinates (Inselberg, 2009). The NEVIS method systematically removes one node from a network and uses PCP to show how removed nodes influence the centrality (measure of how connected a node is to the network) of remaining nodes. Singer and Greenshpan (2009) introduce the terms stability (a node that is not effected by removal of other nodes), topological significance (a node that when removed significantly effects the centrality of other nodes in the network), and backup (alternate paths bypassing a removed node) to describe node functions in a network. They also qualitatively explain how to visually identify these functions on the PCP. The qualitative descriptions require manual identification of key nodes and do not allow an analyst to rank or compare the relative importance of nodes nor work with a large number of nodes.

PCPs can help visualize large-dimensional problems but have also seen limited use because (1) they often have many lines and become busy and crowded (Edsall, 2003), (2) the ordering of axes affects the interpretation of results (Edsall, 2003, Huh and Park, 2008, Albazzaz et al., 2005), and (3) they take time to manually construct (Albazzaz et al., 2005). To address these issues, we introduce a new automated parallel coordinate method and tool to visualize large water resources networks and identify and rank nodes that are (1) stable (their connectivity in the network does not depends on the existence of particular nodes), (2) topologically significant (when removed or added to the network, they cause other nodes to be unstable), and (3) redundant with other nodes. Ranking nodes focuses attention on a few traces in the PCP. The analysis controls for the ordering of the PCP axes, scales to very large networks, and identifies the most promising nodes to subsequently focus computationally-intensive simulation and sensitivity analysis efforts. We illustrate the approach to identify promising areas for agriculture-urban water transfers and other water management actions in the Lower Bear River basin. Below, we describe steps of the tool and its application to the lower Bear River system.

**Methods**

We create an algorithm that uses the NEVIS method (Singer and Greenshpan 2009) to automatically generate the PCP. We then extend the method to identify and rank stable, topologically significant, and redundant nodes. To rank nodes, we develop performance metrics, manipulate the connectivity matrix that describes the network topology, and use summary statistics to evaluate the metrics. In this section, we introduce the tool, define the terms, concepts, and performance metrics on which the tool is built, and then describe how the tool works.

The analysis tool is created using HydroPlatform (Harou et al., 2010) and Microsoft Excel 2007. HydroPlatform is a generic, geographically-based, open-source software platform that can be used with other water resources modeling software to support decision making. The user draws the nodes and links representing the water resources network in HydroPlatform then outputs the network connectivity data as a comma separated value (CSV) matrix that is imported into Excel. The tool uses Excel’s Visual Basic for Applications (VBA) macro programming capabilities to operate on the matrix for multistep calculations and to automate the analysis tool.

We create and build upon terminology introduced by Singer and Greenshpan (2009) for the NEVIS method (Table 1).

Table 1. Terms and definitions of parallel coordinate vocabulary for network analysis.

|  |  |
| --- | --- |
| Term | Definition |
| Adjacency Matrix | A two-dimensional square matrix whose columns and rows enumerate the network nodes. An element in this matrix with a value of one indicates the node represented by the corresponding row is connected by one link to the node indicated by the column. |
| Connectivity Matrix | Square matrix identical in structure to the adjacency matrix where an element value indicates the minimum number of links between the pair of nodes specified by the row and column. Adjacent nodes have a connectivity value of 1. |
| Extracted Network | A network with one node removed. |
| Centrality Value | Measure of how connected a node is to other nodes in the network. Centrality is the inverse of connectivity and higher centrality values indicate the node is more connected. |
| Extracted Centrality | Centrality of a node in an extracted network where another node is removed. |
| Stability | Measure of how much the extracted centrality value for a node changes across extracted networks. |
| Topological Significance | Measure of how extracting a node affects the stability of other nodes in the network. |
| Redundancy | Measure of connection similarity between a pair of nodes. |

The tool has four sequential parts: calculate extracted centrality, create the PCP, calculate pairwise differences, and calculate node stability, topological significance, and redundancy performance indicators. We describe the calculation steps and present some of the equations used in Excel and VBA.

**Step 1: Calculate Extracted Centrality.** First, a user draws the network as a directed graph of nodes and links in HydroPlatform. Then the user exports the square adjacency matrix that identifies the pairs of nodes with unidirectional connections that are one link apart. To represent bidirectional flows as required by NEVIS, we mirror non-zero entries in the adjacency matrix across the primary diagonal (e.g., if node A links to node C, then we also set the value for C to A in the matrix to 1). We then remove one node from the adjacency matrix to create an extracted network. We pivot across rows and down the columns of the extracted adjacency matrix, transition at element values of 1, calculate the smallest number of pivots (links) between each pair of nodes, and use these pivoting results to populate a corresponding extracted connectivity matrix. We calculate each node’s centrality, χ, in the extracted network by summing the inverses of the extracted connectivity values associated with all other nodes (a row of values in the extracted connectivity matrix; Equation 1).

(Eqn. 1)

Where: χi,k = centrality value for node *i* (i=1 to n) in the network with node *k* extracted (k=1 to n; k≠i ); n = number of nodes in the network; j,k = duplicate listings of nodes (j,k=1 to n) that represent, respectively, the columns in the connectivity matrix (j) and extracted network (k); and distance (i, j, k) = value representing the minimum number of links between nodes *i* and *j* in the network with node *k* extracted*.*

Calculating Eqn. 1 for each node *i* in a particular extracted network generates a column vector of extracted centrality values for that extracted network. We repeat the extraction, distance, and centrality calculations for each node and generate an extracted centrality matrix whose columns represent the extracted nodes, rows represent the node affected by the extraction, and primary diagonal elements are undefined. A row in this extracted centrality matrix shows how node extractions affect the centrality of a particular node.

**Step 2: Create the Parallel Coordinate Plot.** We then plot the extracted centrality matrix in parallel coordinates as suggested by Singer and Greenshpan (2009). We plot centrality values as ordinates, list extracted networks as abscissa, and plot each node’s row-vector of centrality values as a horizontal trace spanning the abscissa. The abscissa represents the extracted networks and simultaneously shows on the same plot the effects of each node extraction. Together, listing the extracted networks along the abscissa comprise parallel axes along which centrality values scale from zero to the largest value observed. The result is a parallel coordinate plot. For example, a network with 10 nodes requires 10 axes and 10 traces spanning the axes to capture the 36 node interactions (how removing one node affects the others). From the PCP, we visually identify unstable nodes as nodes whose traces have vertical drops. Topologically significant nodes are locations on the x-axis where extracting a node causes multiple traces to drop.

**Step 3: Calculate Pairwise Differences.** To quantify node stability, topological significance, and redundancy shown on the PCP and to quantify these characteristics independent of the order in which the extracted networks are plotted on the PCP, we calculate the difference between each pair of extracted network centrality values along a trace and call these differences pairwise differences. For a particular node and trace, there are (n-1)+(n-2)+(n-3)+…+(2) = [(n-1)(n-2)/2] pairwise differences to consider and these differences allow us to consider all combinations and interactions across extracted network axes and move beyond analysis for the particular ordering of the extracted network axes shown on the PCP (i.e., the axes immediately to the left or right of a selected axis). We tabulate a histogram of pairwise differences for each node across difference intervals the user provides. Here, frequency is the count of the number of pairwise difference values within a difference interval divided by the [(n-1)(n-2)/2] total number of pairwise differences for a node. We classify candidate redundant nodes as nodes that have similar histograms (they also have similar horizontal traces on the PCP).

**Step 4: Calculate Performance Indicators.** To quantify a node’s stability, we average all the pairwise differences associated with a trace. Higher average pairwise differences indicate nodes whose traces have larger drops. The most unstable nodes have the highest average pairwise differences.

We quantify topological significance by examining two factors associated with the traces of extracted centrality values: the number of drops at an extracted network axis and the magnitude of each drop. Multiple traces that drop at the same extracted network indicate that extracting that node causes many nodes to be unstable. Extracted nodes that cause large numbers of traces to drop and large magnitude drops are topologically significant. We establish a minimum drop threshold to count the number of drops that an extracted node causes and measure the magnitude of each drop. We use the average of the pairwise differences for all traces at an extracted node to quantify the magnitude of drop an extracted node causes. We rank each node for number and magnitude of drops. The node having the highest average number and magnitude of drops is the most topologically significant.

To determine redundancy, we first identify candidate redundant node pairs that have pairwise difference histograms that are less than 0.5% different for each delineation. We then compare the row vectors from the connectivity matrix for each node in the pair. We quantify redundancy as a percentage by comparing the number of common connectivity values to the maximum number of connections (Equation 2).

(Eqn. 2)

Where: Ri,j = redundancy between nodes *i* and *j* expressed as a percent; ci,j = count of number of nodes for which nodes *i* and *j* have the same connectivity value (i.e., Ci,l = Cj,l where C represents the connectivity matrix and *l* is another node in the network l≠i≠j); and n-2 = maximum number of common connections a node pair can share, the number of nodes minus 2 because each node in the pair cannot connect to itself.

To automate calculating the extracted centralities, plotting in parallel coordinates, calculating pairwise differences, and calculating network performance indicators, the user must provide three inputs: adjacency matrix from HydroPlatform, the PC drop threshold to determine topological significance, and intervals used to generate the histograms of pairwise differences. The tool outputs the PCP, ranks each node’s stability and topological significance, and lists node pairs that are redundant.

**Applications**

We demonstrate the tool for two small illustrated networks. Then, we apply the tool to inform management of the much larger Bear River water system that has 56 nodes within Idaho and Utah.

**Illustrative Networks.** We present illustrative (i) clique and (ii) hub and spoke networks that are simple and uniform in construction but with very different structure to introduce how the tool works, verify the tool outputs are correct, and highlight relationships among the performance indicators that quantify stability, topological significance, and redundancy. We describe each network and present key results.

Clique. In a clique network, each node is adjacent to all other nodes (Figure 2-A1). The adjacency matrix for the clique network has all ones (except on the primary diagonal) and running our tool yields a PCP where all of the traces have straight horizontal traces (Figure 2-A2). There are no drops in the traces across the extracted node axes so all the nodes are stable and none of the nodes are topologically significant. Similarly, each row of the connectivity matrix has n-1 connections each with a connectivity value of 1 and each pair of nodes has the same number of connections. All nodes are redundant with an R = 100% for each node pair. Together, the results confirm what is apparent from visual inspection that removing a node from a clique network does not affect the connectivity or centrality of other nodes.

Hub and Spoke. In a hub and spoke network, a single hub node (A) is the sole connection for all exterior spoke nodes (B to J in Figure 2-B1). The adjacency matrix for this network has all ones in the row for the hub node. In the connectivity matrix, each spoke is two links away from each other spoke and the hub is adjacent and one link from all other nodes. The PCP shows that the centrality values for all spoke traces drop at the axis A representing extraction of the hub node whereas the trace for the hub is a straight horizontal line and does not drop when any spoke is extracted (dashed blue line in Figure 2-B2). Since the traces for the spoke nodes drop, the spokes are unstable. Drops occur when the hub node is extracted and make the hub topologically significant. The traces for all the spokes are identical and make the spoke nodes redundant with each other with an R = 100%. Again, the results confirm what is apparent from visual inspection that removing the hub node from a hub and spoke network affects the network connectivity.

|  |  |  |
| --- | --- | --- |
| Network Characteristic | A. Clique | B. Hub & Spoke |
| 1. Directed Graph |  |  |
| 2. Parallel Coordinate Plot |  |  |
| 3. Stable | All | Hub (A) |
| 4. Topologically Significant | None | Hub (A) |
| 5. Redundant | All | All spokes (B-J) |

Figure 2. Parallel coordinate analysis of illustrative networks.

These illustrative networks provide a way to verify the accuracy of our network analysis tool and visualize and quantify stable, topologically significant, and redundant nodes in simple and uniform networks where we might otherwise be able to identify the key nodes absent the tool. The results also show that the performance metrics of stability, topological significance, and redundancy are not mutually exclusive – for example nodes can be both unstable and redundant (spokes in a Hub and Spoke network) or redundant but neither topologically significant nor unstable (all nodes in a Clique network). We now apply the tool to the much larger and irregular lower Bear River system to identify promising locations for agricultural to urban water transfers and water management actions.

**Bear River Network.** The Bear River watershed comprises 7,500 square miles of agricultural, urban, federal, and state lands in southeastern Idaho, northeastern Utah, and southwestern Wyoming and is the largest tributary of the Great Salt Lake with an average annual inflow of 1.2 million acre feet (Mesner and Horsburgh 2012). The primary water uses in the basin are for agriculture, municipal, industrial, power generation, recreation, and the environment.

The lower Bear River system used in this analysis stretches from Southeastern Idaho to the Great Salt Lake, Utah. The Utah Division of Water Resources developed a Bear River simulation model to look at water system sustainability over a 50-year historical record and plan for future infrastructure (new dams and diversions) and operational changes (reservoir reoperations, impacts of water conservation, and reallocations among demand sites; Utah Division of Water Resources, 2004). The model represents the system with 56 nodes (9 reservoirs [5 existing and 4 proposed], 12 municipal, agricultural, and environmental service areas, 13 flow junctions) and 74 linkages (Figure 3, Utah Division of Water Resources, 2010). One environmental wetland service area is the Bear River Migratory Bird Refuge while other incidental riparian water uses occur at various junctions along the main stem of the Bear River.

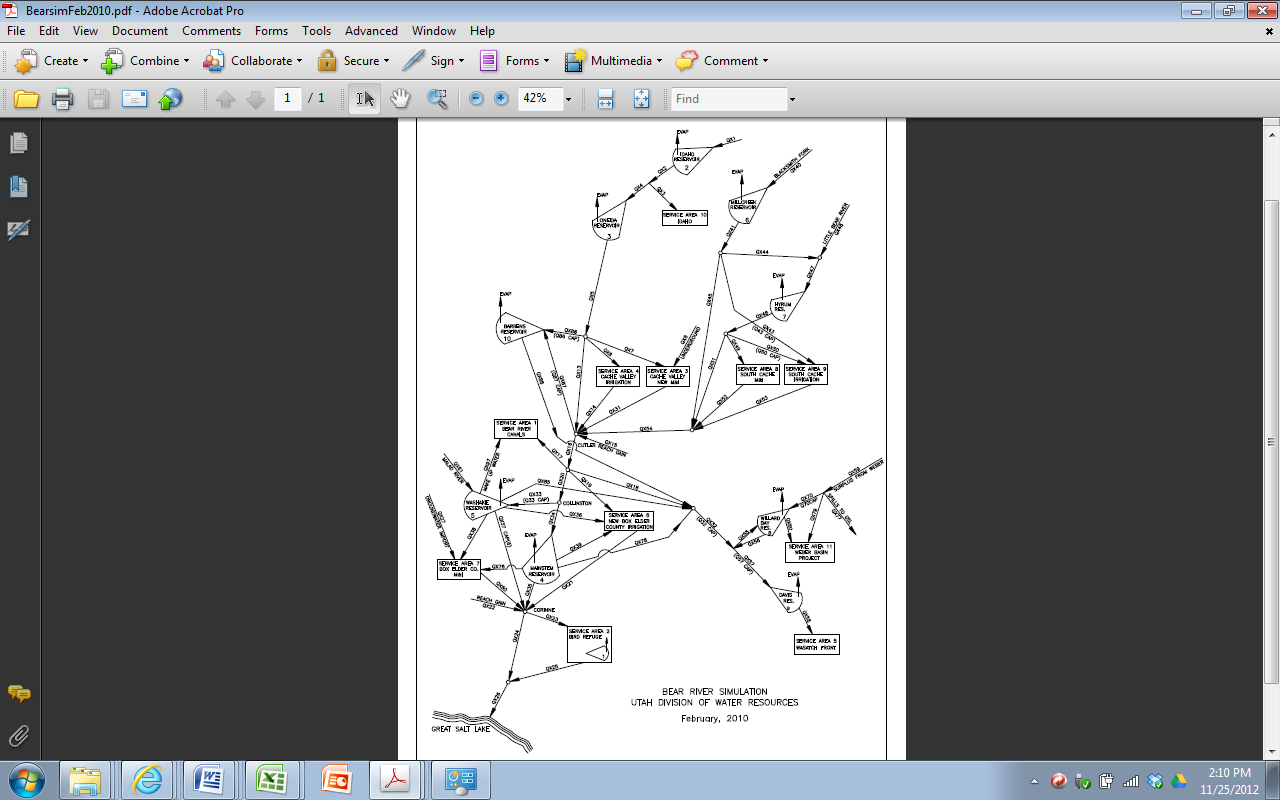


Figure 3. Bear River network used for analysis.

Utah’s urban population continues to grow along the Wasatch front in Salt Lake, Davis, and Weber counties and planners project the need to transport Bear River water to these areas (Mesner and Horsburgh 2012). Water transfers from agricultural to urban uses could change how the water system functions. From what agricultural areas should managers transfer water to meet future urban demands? Additionally, where might it be appropriate to remove or build dams, implement conservation measures, develop new local resources, monitor flows, or protect environmental and ecosystem services?

We drew the Bear River network schematic in Figure 3 as a directed graph in HydroPlatform. We input the adjacency matrix into Excel and steps 1 and 2 of the tool produce the PCP. On the PCP (Figure 4), extracting junctions cause multiple traces to drop considerably whereas traces are relatively flat (few drops) when extracting individual water sources and sinks. These latter source and sink nodes thus have little topological significance.

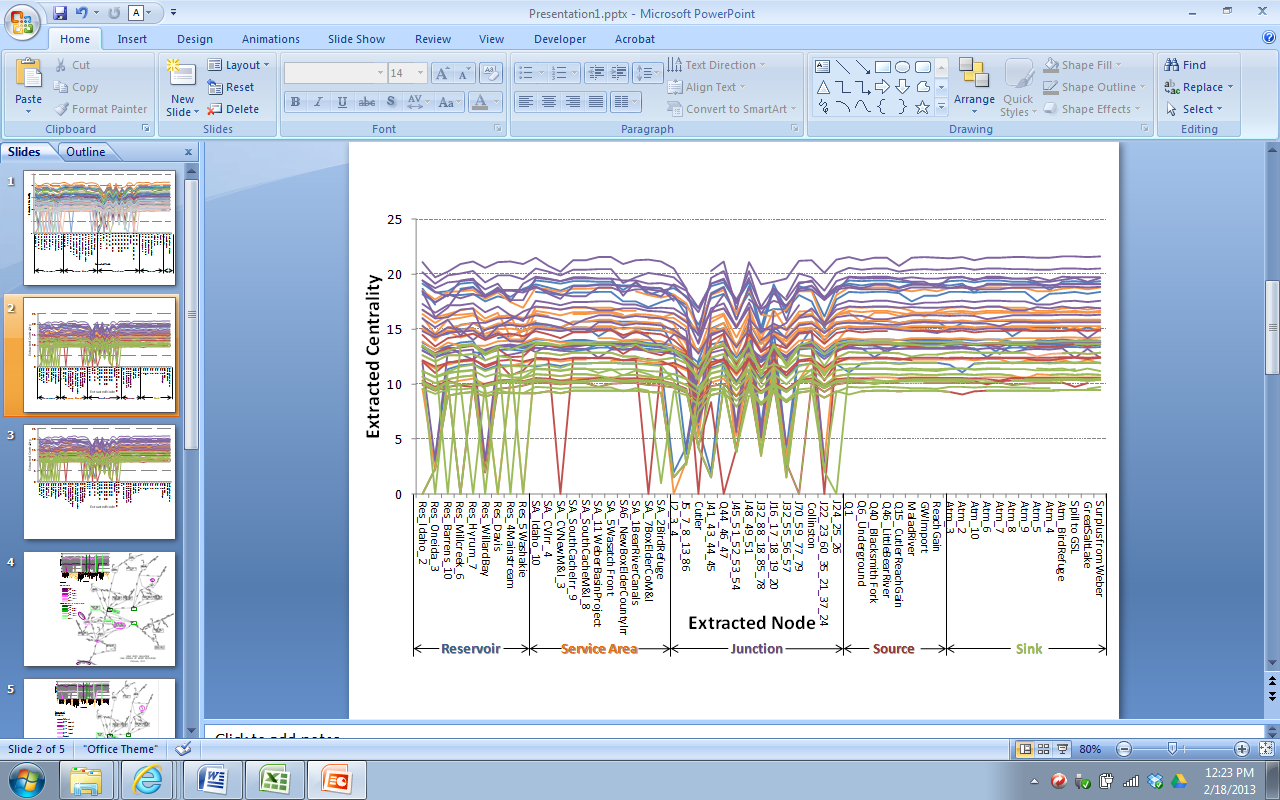


Figure 4. Parallel coordinate plot for Bear River network.

We then identified and ranked each node’s stability, topological significance, and redundancy. The most unstable nodes—the sources and sinks the Great Salt Lake, Malad River, and Evaporation from Hyrum Reservoir—have traces with many drops and are connected to only one other node (Figure 5 and Table 2). The three most topologically significant nodes—the junctions Cutler, J45-51, and J32-88—have many connections and serve as conduits to connect large portions of the network to the main stem of the network (Figure 5 and Table 2).

Table 2. Most and least stable and topologically significant nodes in the Bear River network.

|  |  |  |
| --- | --- | --- |
| Rank | Stability | Topological Significance |
| 1 | Washakie Reservoir | Cutler Reservoir |
| 2 | Junction 22-60 | Junction 45-51 |
| 3 | SA7 Box Elder M&I Users | Junction 32-88 |
| 54 | Evaporation from Hyrum Reservoir | Q15 Cutler Reach Gain |
| 55 | Q61 Malad River | Collinston |
| 56 | Great Salt Lake | SA1 Bear River Canals |

The topologically significant junctions labeled Cutler, J45-51, and J32-88 have the most connections with 7, 5, and 5 links to other nodes, respectively. The Corrine junction also has seven connections but is not topologically significant because it does not serve to connect a group of nodes to the main network as do the three most topologically significant nodes.

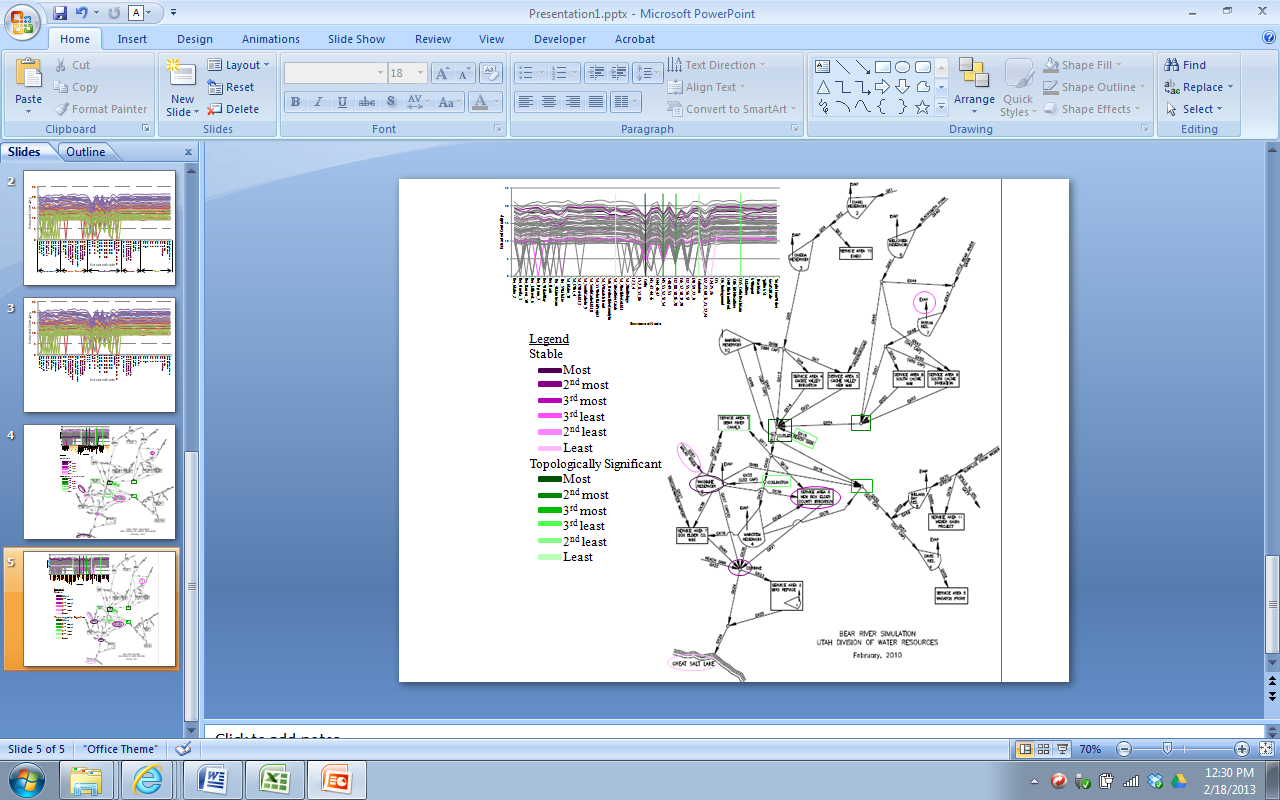


Figure 5. Location and ranking of the stable and topological significant nodes in the Bear River network.

Five node pairs have redundancy values of more than 96% and the first two pairs with 100% redundancy are connected to the same and only one node (Table 3). The small differences between the nodes in the pairs with redundancies of 98% and 96 % (SA4 Cache Valley Irrigation and SA3 Cache Valley New M&I; and Idaho Reservoir and SA10 Idaho) are that one node in each pair has one different connection. SA 4 Cache Valley Irrigation and SA3 Cache Valley New M&I are both connected to J5-86 and Cutler; Cache Valley New M&I is also connected to Q6.

The SA11 Weber Basin Project and J70-79 node pair has three different connections, but the two nodes are connected to each other.

Table 3. Five most redundant node pairs in the lower Bear River network.

|  |  |  |
| --- | --- | --- |
| Node 1 | Node 2 | Redundancy |
| Spill to GSL | Surplus from Weber | 100% |
| Malad River | Evaporation from Washakie Reservoir | 100% |
| SA4 Cache Valley Irrigation | SA3 Cache Valley New M&I | 98% |
| Idaho Reservoir | SA10 Idaho | 96% |
| SA11 Weber Basin Project | Junction 70 and 79 | 96% |

**Discussion**

Our automated PC tool can identify and rank the stable, topologically significant, and redundant nodes in a network. In the lower Bear River network, the least stable nodes are the Great Salt Lake, Malad River, and Evaporation from Hyrum Reservoir while the most topologically significant nodes are Cutler, Junction 45-51, and Junction 32-88. Five node pairs have over 96% redundancy. Nodes that are connected to only one node are more likely to be unstable, as is the case for the Bear River network’s three least stable nodes. The most topologically significant nodes are junctions that serve as conduits that connect large branches of the network to the rest of the network. Junction 45-51 connects the south Cache Valley nodes to the network and Junction 32-88 connects the Weber area, to the rest of the network. Several of these topologically significant locations are on primary tributaries just before the tributary confluences with the main stem. The node centrality and topologically significance metrics are a way to quantify and rank this connectivity.

The service areas of Cache Valley Irrigation and Cache Valley New M&I are redundant and provide the opportunity to reroute water in the event of delivery system failure at one node. Nodes with more connections are less likely to be redundant. Redundancy can be both positive and negative from a management standpoint. Redundancy adds operational flexibility but also increases costs to create the redundancy.

Removing nodes that are redundant or not topologically significant will have little effect on network connectivity. Having a redundant pair of nodes means that absent one node, managers can still route water through the other node. Conversely, losing a topologically significant node will affect many other nodes and may mean water cannot be routed to the desired destination.

**Implications for Potential Water Transfers.** Here, we use the network analysis results for the lower Bear River system to analyze the impact of transferring water from agriculture to urban users and identify promising sources of transfers. Results show the Cache Valley Irrigation service area is 98% redundant with another node and suggest that Cache Valley Irrigation has little influence on network connectivity. Its topological significance rank is 53 out of 56 and further emphasizes the redundancy result. The New Box Elder County Irrigation and South Cache Irrigation service area nodes have topological significance ranks of 49 and 51, respectively, and indicate that these service areas also have little effect on the network. Together, the analysis suggests the Cache Valley, South Cache, and New Box Elder County Irrigation service areas may be good candidates for sources of agricultural to urban transfers if the goal of the transfer is to leave intact the connectivity of the remaining parts of the Bear River water system. The effects of using these suggested sources for water transfers need to be further tested and explored with more detailed simulation modeling.

**Implications for Water Management.** In addition to identifying promising sources of water transfers, the PCP tool can inform many other water management decisions. For example, we can use the tool to identify existing or proposed dams to remove, build, or not build. In the Bear River network, Hyrum Reservoir has a topological significance rank of 39 out of 56 which is the lowest of all the reservoirs and suggests that removing this dam would have little effect on network connectivity. Oneida and Willard Bay reservoirs have high topological significance values and indicate their removals would significantly impact service areas and other nodes in the Bear River network (Table 4). Additionally, the proposed Washakie, Mainstem, and Millcreek reservoirs also have high topological significance values which suggest it important to build these reservoirs.

Managers should consider backup water conservation measures, development of new alternative supplies, and flow monitoring at unstable nodes because water supply to these locations is easily affected by changes at other network nodes. In the Bear River system, the least stable service areas are the Wasatch Front, Idaho, and Weber Basin Project. These service areas would benefit most from conservation and alternate supplies for times when issues occur such as low surface water availability, reservoir storage, or breaks in water transmission lines. Interestingly, Weber Basin Water Conservancy District managers, who oversee the Weber Basin Project, have been steadily promoting water conservation programs over the last decade. Additionally, managers should also monitor water flows at unstable nodes like the Great Salt Lake and Malad River because these nodes can see variable flows subject to activities at many other nodes in the network.

Managers should also protect areas with high topological significance if they provide environmental service because degradation or removal of these areas will both negatively impact ecosystem services at that location as well as the overall ecosystem and low stability. The Bear River environmental areas J5-86 and J22-60 have high topological significance and need to be protected.

Together, stability, topological significance, and redundancy results can help managers identify agricultural areas for potential water transfers to urban use and dam removal sites. The performance metrics can also identify areas needing water conservation measures, new alternative supplies, monitoring, or environmental protection.

Table 4. Bear River network nodes of interest for water management.

|  |  |  |  |
| --- | --- | --- | --- |
| Node Type | Node Name | Stability | Topological Significance |
| Reservoirs | Washakie | 1 | 11 |
| Mainstem | 8 | 12 |
| Hyrum | 15 | 39 |
| Oneida | 22 | 7 |
| Willard Bay | 30 | 8 |
| Millcreek | 33 | 13 |
| Service Areas | Weber Basin Project | 36 | 47 |
| Idaho | 37 | 46 |
| Wasatch Front | 38 | 48 |
| Environmental | Junction 22-60 | 2 | 5 |
| Junction 5-86 | 11 | 6 |
| Bird Refuge | 16 | 17 |

**Future Work.** With our PCP tool, the stability, topological significance, and redundancy of network nodes are determined solely based on a binary (yes/no) and bi-directional definition of the connectivity of adjacent nodes. We also assume that each node is of equal importance, do not weight one node over another, nor limit the magnitude of flows through links. For example, the analysis assumes reservoir evaporation has the potential to have the same influence on the network as the inflow to or release of water from the reservoir to the next downstream node. Similarly, when two nodes are adjacent, the analysis allows unlimited movement between the nodes in either direction.

Further analysis should include the flow direction and magnitude such as each link’s flow capacity (volume or flow rate). When considering these factors, nodes may no longer be redundant if water cannot spill from a reservoir or a capacity constraint or other consideration limits flow through a link. We can incorporate flow magnitude by including into the calculation of extracted centrality a factor that represents the minimum or maximum flow capacity, average flow, or another quantifiable attribute that describes the pathway of links that connect each pair of nodes. For example, if flow capacity is the relevant link attribute, two nodes have a distance of three, and the capacities of the three intermediary links between the nodes are 10, 5, and 7 cms, then 5 cms would be the capacity associated with the pathway between the two nodes and the value to use along with the distance of three to calculate centrality. Thus, centrality becomes a function of both the distance between nodes and the attribute value associated with the path of links connecting the node pair (Equation 3).

(Eqn. 3)

Where: χ’i,k = modified centrality value for node *i* (i=1 to n) that considers both distance and a pathway attribute in the network with node *k* extracted (k=1 to n; k≠i ); g = function to weight and combine the distance and the pathway attribute values when calculating χi,k; and vi, j, k = value of the attribute (from 0 to infinity) associated with the path of links between nodes *i* and *j* in the network with node *k* extracted. *di,j,k*, and *i*, *j*, and *k* are defined as previously*.*

We can include flow direction in the analysis by explicitly considering direction in the definitions of adjacency and connectivity. We can include direction by not mirroring the non-zero elements across the primary diagonal of the adjacency matrix (e.g., downstream nodes link to upstream but not the reverse). For example, Cutler is adjacent to SA4 Cache Valley Irrigation node and five other upstream nodes but not adjacent to its downstream junction. However, the downstream junction is adjacent to Cutler. When including flow direction, we calculate the stability, topological significance, and redundancy by drawing the network with links defined from the destination node to the source node (flow from A to B is represented as B to A). The link direction is represented as the reverse of the flow direction because downstream nodes rely on upstream nodes for water supply. Nodes drawing water supply from one or a few upstream sources are more likely to be unstable and their topological significance will likewise depend on the stability of the connected downstream nodes when the node is extracted from the network. A source node (with no upstream water sources or links) will not have a horizontal trace on the PCP. Its stability is undefined but its topological significance will depend on the stability of the downstream nodes connected to the source when the source node is extracted from the network. Together, improvements like adding flow direction and magnitude will likely influence the classification and ranking of stable, topologically significant, and redundant nodes and represent important areas of further work.

**Conclusions**

Modeling complex water resources networks requires significant computational effort. Networks have many nodes and multiple decision variables which are difficult to visualize in Cartesian coordinates. These difficulties are exacerbated when removing individual nodes to study effects on other nodes in the system. We developed an automated parallel coordinate plotting and analysis tool to identify key network nodes that are (1) stable (their roles do not depend on the existence of particular nodes), (2) topologically significant (when removed or added to the network, these nodes cause other nodes to be unstable), and (3) redundant (node pairs that have similar connections). We applied the tool to identify critical agricultural to urban water transfers in the 56-node, lower Bear River water system that extends from southern Idaho to the Great Salt Lake, Utah. Nodes representing the Great Salt Lake, Malad River, and Evaporation from Hyrum Reservoir are connected to only one other node and are the least stable. The three most topologically significant nodes that cause other nodes to be unstable are the junctions Cutler, J45-51, and J32-88. The Cache Valley Irrigation and New M&I service areas are one of five node pairs with over 96% of the same connections. The redundant Cache Valley Irrigation service area and low topological significance of New Box Elder County Irrigation and South Cache Irrigation service areas identify these sites as promising sources of water transfers from agriculture to urban uses since they have little effect on the Bear River network connectivity. These three areas merit further investigation with more detailed simulation modeling. Our tool can also help inform water system planning and management such as identify candidate sites for dam removals, conservation measures, development of new alternative supplies, monitoring, or environmental projection. The PCP and network analysis tool scales to very large networks and identifies the key nodes on which to concentrate further time-intensive modeling and sensitivity analysis.

**References**

Albazzaz, H., Wang, X.Y., Marhoon, F. (2005). “Multidimensional visualization for process historical data analysis: a comparative study with multivariate statistical process control.” *J. of Process Control,* 15(3), 285-294.

Andrienko, G. and Andrienko, N. (2001). “Exploring spatial data with dominant attribute map and parallel coordinates.” *Computers, Environment and Urban Systems,* 25(1): 5-15.

Choi, H., Heejo, H., Hyogon, K. (2009). “Fast detection and visualization of network attacks on parallel coordinates.” *Computers & Security,* 28(1): 276-288.

Edsall. R.M. (2003). “The parallel coordinates plot in action: design and use for geographic visualization.” *Computational Statistics & Data Analysis*, 43(4): 605-619.

Harou, J., Pinte, D., Tilmant, A., Rosenberg, D. E., Rheinheimer, D., Hansen, K., Reed, P., Reynaud, Medellin-Azuara, J., Pulido-Velazquez, M., Matrosov, Padula, and Zhu, T. (2010). "An open-source model platform for water management that links models to a generic user-interface and data-manager." 2010 International Congress on Environmental Modelling and Software 'Modelling for Environment’s Sake', Fifth Biennial Meeting, Ottawa, Canada.

Inselberg, A. (1985). “The plane with parallel coordinates.” *The Visual Computer* 1, 69-91.

Loucks, D.P., van Beek, E., Stedinger, J.R., Dijkman, J.P.M., Villars, M.T. (2005). Water resource systems modeling: Its role in planning and management. *Water Resources Planning and Management: An Introduction to Methods, Models and Applications.*UNESCO, Paris, France.

Matrosov, E.S., Harou, J.J., Loucks, D.P. (2011). “A computationally efficient open-source water resource system simulator – Application to London and the Thames Basin.” *Environmental Modelling & Software*. 26(2011) 1599-1610.

Mesner, N. and Horsburgh, J. (2012). “Watershed Description.” *Bear River Watershed Information System,* < <http://www.bearriverinfo.org>> (Nov. 25, 2012).

Rosenberg, D.E. (2012). Visualizing high dimensional data. CEE 6930 Hydroinfomatics course materials. <https://usu.instructure.com/courses/127332/wiki/lecture-materials> (Dec. 12, 2012).

Singer, Y. and Greenshpan, O. (2009). Network visualization and analysis with parallel coordinates. *Parallel coordinates: visual multidimensional geometry and its applications*, A. Inselberg, ed., Springer Dordrecht, New York, N.Y., 460-475.

Utah Division of Water Resources. (2010). Map of Bear River simulation area.

Utah Division of Water Rights. (2004). Bear River Basin: planning for the future. <<http://www.water.utah.gov/planning/SWP/bear/bearRiver-1A.pdf>> (Dec. 14, 2012).

Wegman, E.J. (1990). “Hyperdimensional data analysis using parallel coordinates.” *J. of the American Statistical Association,* 85(411): 664-675.