Leah Meeks

Civil and Environmental Engineering

Utah State University

Logan, UT 84322

Phone: (435)881-8697

E-mail: l.meeks@aggiemail.usu.edu

August 19, 2014

**SUBJECT: Re-submission of Manuscript #WRENG-1239 to *Journal of Water Resources Planning and Management***

Candice Gooch

Editorial Coordinator

JWRPM

Dear Ms. Gooch:

“Identifying and ranking stability, topological significance, and redundancies in water resource networks to guide detailed simulation modeling” by Leah Meeks and David E. Rosenberg is being submitted the manuscript for publication in the *Journal Water Resources Planning and Management*. It is a revised version of the original manuscript #WRENG-1239 “Identifying stability, topological significance, and redundancies in water resource networks using parallel coordinate plotting.”

A summary of changes with responses to the original manuscript #WRENG-1239 reviewers’ comments is attached to this letter. Should you need to contact me, please use the above e-mail address or phone number.

Sincerely,

Leah Meeks

Corresponding Author

Enclosures

Summary of Responses to Reviewers Comments

Publication sizing guide

Manuscript

List of Figure Captions

Figures: 1a and 1b, 2, 3, 4, and 5

Copyright Transfer Agreement

**Summary of Changes:**

We appreciate the time the reviewers and associate editor have taken to provide extremely helpful feedback to improve the content, writing, and contribution of the paper. We have revised with the following principal changes made largely in response to the reviewer comments:

1. Added flow direction to the analysis
2. Clarified in the abstract and introduction the study objectives which are to quantify the performance metrics, rank nodes according to those metrics, apply the results to water resources to inform management decisions, and automate the calculations so that they can be used to identify nodes to subsequently focus computationally-intensive simulation and sensitivity analysis
3. Better articulated the difference between previous work of NEVIS, parallel coordinate plots, and our contributions which are to (i) quantify, automate, and rank nodes in a network by the performance metrics of node stability, topological significance, and redundancy, (ii) include flow direction, and (iii) apply the ranked metrics to inform management of a water resources network
4. Edited the manuscript to make it more clear and consistent

Below, editor and reviewer comments appear in plain black text and our response is in *italicized red*.

Editor Comments:  
  
Editor: Based on the reviews, it is recommended that the author should revise and resubmit the manuscript. The author is encouraged to review the past JWRM publications on this subject and to take the reviewer comments into consideration in improving the paper. However, please submit a list of changes or a rebuttal against each point raised by the reviewer with your revised manuscript. Please note that the earlier we receive your revised manuscript, the earlier we can process it. Thanks for your interest in the Journal of Water Resources Planning & Management. We look forward to receiving the revised manuscript from you. The reviewer comments are listed below.  
  
  
Editor: See comments below. Major revisions are warrant. The decision on publication of this paper is thus deferred until the authors have a chance to revise and resubmit the paper. The revisions should address the concerns and suggestions made by the reviewers. A detailed description of how those concerns have been addressed in the revision should also be submitted with the revised paper. The authors should pay careful attention to the length of their revised paper and ensure that it complies with the ASCE guidelines for length of Technical Papers and submit a filled out sizing guide spreadsheet with the revised paper. The authors are requested to review the recent research literature (including web published In Press and Posted Ahead of Print papers) and include them in their revised submission.

The sizing guide spreadsheet has been completed (estimating 7 pages) and submitted.

We added a description of UCINET, a social network analysis tool, to the literature review in the introduction. Some of its methods (drawing a directed graph of the network) were used in creating the tool presented in the manuscript. In general, we expanded the literature review section for the citations from the first draft of the manuscript to more clearly explain what prior work did and how the current paper expands on that prior work.

Associate Editor: The manuscript entitled "Identifying stability, topological significance, and redundancies in water resource networks using parallel coordinate plotting" seeks to exploit parallel axis plotting to inform effects of changes to water resources network structure and management. Three disparate reviews have been received ranging from Reject to Accept As Is.  I would discount the Accept As Is review as it has no technical feedback and does not appear to have carefully reviewed the work. Reviewers #1 and #3 have provided detailed feedback on the manuscript and agree that the study requires substantial revision before it is worthy of publication.  In my own review of the work, the results are very preliminary and are limited in their generalizable contributions.  The visualization technique and tools are discussed in a tutorial fashion with limited insights in the actual application.  The paper has a broad number of caveats that raise the concern that the work is too preliminary.  I would recommend that the authors strongly revise their study using the reviewer feedback.  The revised manuscript needs to strongly improve the clarity and generality of its contributions.

We have expanded the scope of the work and heavily edited the manuscript to address the reviewer comments. Many of the tutorial-like calculation details have been removed. We retain the reasoning behind the calculations (e.g., pairwise differences between nodes on each axis to remove the bias of axes ordering) to explain the logic used to create the performance metrics and the tool as well as emphasize the contribution beyond prior work. The main contributions of this method and tool are to identify key network nodes through quantification and ranking of stability, topological significance, and redundancies for subsequent focus with more detailed simulation modeling. Additional contributions include considering uni-directional flows across links such as typically occurs in a water resources network and controlling for the ordering of axes on the parallel coordinate plot. The discussion section also now highlights how the ranked node stability, topological significance, and redundancy results can be used to inform the management of the network.

Reviewer #1: The paper by Meeks and Rosenberg seeks to demonstrate a parallel coordinate plot (PCP) approach for diagnostics of a water supply network.  The paper is of interest to the readership base, and the tool could be useful to real water managers.  The syntax of the paper needs some revision, and in the revised manuscript the authors should ensure that all the implications of their work are clearly discussed and presented.  I am happy to recommend publication subject to the following general and specific comments.  
  
General Comments:  
  
1. There is too much use of the first person ("We...") in the manuscript.  In my experience, ASCE style prohibits any use of the first person at all -- so since the authors will have to change this eventually they might as well do it in the revision.

We have generally removed first person from the manuscript.

2. In the beginning of the paper there is a lot of detail on the calculations being performed in the tool.  Later, in the Discussion, there is good insight on the relevance to water management.  I suggest eliminating some of the details of the generic method in order to write more about the implications of the approach.  For example, the authors mention that an improvement to the method would include flow direction and magnitude.  Does this mean that the method can be plausibly used "as-is", or does this improvement really need to be added before policy-relevant information can be taken from the tool?

Thanks for this observation. We have removed many of the calculation details such as equations and how Excel VBA steps though calculations for each node. We retain the reasoning behind the calculations (e.g., pairwise differences between nodes on each axis to remove the bias of axes ordering) to explain the logic used to create the performance metrics (stability, topological significance, and redundancy) and the tool.

The main contribution of this method and tool is to identify key network nodes through quantification and ranking of stability, topological significance, and redundancies that then should be used to focus detailed simulation modeling. The revision now includes flow direction, which is a step further in the analysis so that the method and tool now better represent what is actually happening in water networks. We agree with the reviewer that including flow magnitude will provide additional information but also feel that the method as presented with flow direction provides significant policy relevant information such as now expanded on in the discussion section. The discussion now describes more ways that results that consider flow-direction can be used in policy making.

3. There should be a clear distinction between whether the authors used the existing NEVIS approach or they created their own.  Furthermore, perhaps the authors should provide a citation or link to where users can download a copy of the software for themselves.

Yes, this suggestion is definitely helpful. The manuscript now more clearly delineates in the abstract, introduction, and section 2 between our work and previous work including NEVIS. NEVIS provides a way to calculate the extracted node data as required for and prepare the Parallel Coordinate Plot. NEVIS also qualitatively defines the three performance metrics (node stability, topological significance, and redundancy) for a general network with bi-directional links between nodes. Our work builds upon NEVIS by considering uni-directional links as well as quantifying the performance metrics and ranking nodes. Our quantification and node ranking also address issues raised by other previous work such as bias created by the ordering of axes on the parallel coordinate plot. We then show how the tool outputs can be used to inform water management issues (e.g., water transfers, developing new sources, environmental protection). We have added a link (https://github.com/lmeeks/RANK) in the methods to guide readers to the code repository where they can download the software tool and documentation.

Specific Comments:  
  
1. [line 51] "Loucks" is spelled wrong in the citation.

Thank you; this has been corrected.

2. [line 52] Does the Wegman reference specifically cover network analysis?  It seems like the citation is misplaced.

Thank you, no. The Wegman reference has been moved to the next paragraph discussing the high dimensionality of parallel coordinate plots.

3. [line 56] "the analysis can suggest improvements..." There is a disconnect between suggesting improvements to the modeled network versus the real-life system.  That is, it is exceedingly difficult to remove a large, built reservoir even if the model says so.  However, one could improve the fidelity of the modeled representation (i.e., removing a redundant water source or sink, as suggested in the manuscript) which may or may not correspond to a real node in the actual system.

The wording has been changed to indicate the analysis can guide models that deal with water management rather than water management itself.

Thank you for pointing out the model fidelity contribution of this tool. The discussion section has been edited to include using RANK to improve model representation of the physical system.

4. [line 62-63] "two or three dimensions"  While this statement is true, parallel coordinate plots only show the relationship between two dimensions at a time.  In Figure 1, for example, one can see how X1 relates to X2, but not how X1 relates to X3.  Line color (such as green vs. pink in the figure) and interaction, which allows one to change the axes in real time, can help alleviate these issues.

Yes, visual inspection of parallel coordinates only shows the relationship between two dimensions at a time (axes that are next to each other). We highlighted this issue in the literature review of parallel coordinates. Axes ordering is a shortcoming of NEVIS as it relies on visual inspection to compare nodes and determine if any nodes are unstable, topologically significant, or belong to a redundant node pair. A major contribution of this work is the quantification of performance metrics in a way that controls for the ordering of axes instead of relying in visual inspection. We address the issue of visually comparing X1 and X2 but not X1 and X3 in our analysis by using pairwise differences to compare each pair of axes instead of just the axes arbitrarily next to each other. The text now more clearly explains how these comparisons are done and the reasons for them.

5. The literature review, starting on line 70, could be improved.  For the Albazzaz et al. study, it isn't really relevant that the authors tracked 38 variables for 527 days -- how did the authors do it?  How does this relate to the current study?  Also, I don't see the connection that viewing a large number of dimensions helps users understand their systems better [line 70].  Why?  More dimensions has the potential to be confusing without proper supporting analysis.

The literature review has been expanded to highlight how parallel coordinates have been used in previous work and how that work has influenced or directed our work. The Albazzaz et al. study combined the use of parallel coordinates and statistics to classify data which helped lead us to using statistical analysis (mean, histograms, etc.) on the parallel coordinate plot data for performance metric quantification.

We also point out shortcomings from previous work that we address though our work: (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), (3) they take time to manually construct (Albazzaz et al., 2005), (4) the data analysis has not been automated (Albazzaz et al., 2005), (5) variables on distant axes cannot be visually compared (Edsall, 2003) [see also response to the prior comment], and (6) parallel coordinates analysis currently only allows for qualitative data comparison based on visual inspection. We have also shortened the review of Albazzaz’s work to focus on the key relevant issues for our work.

The wording has been changed: viewing a large number of dimensions on one plot versus looking at multiple Cartesian plots.

6.  As per my general comment 3 above, there should be a clear distinction between NEVIS and the new method.  In the paragraph starting on line 89, the authors mention some criticisms of NEVIS.  Then, starting on line 117 they say they use the NEVIS method.

Yes, this is a great suggestion, and our response to the prior comment #3 applies here as well. The introduction now more clearly explains the NEVIS work, shortcomings, as well as our improvements.

7.  It is appropriate to show the first three terms in Table 1 on line 135.  However, the definitions of the last 5 terms in the table are confusing; the terms are better explained in the supporting text that comes after.  I think the readers may get confused by the short definition of, say, Centrality Value here, before they read it in the text later.  Some change in the organization may be warranted.

Table 1 now has definitions for adjacency matrix, connectivity matrix, extracted network, and the key node attributes of stability, topological significance, and redundancy. As readers go through the paper, the table provides a quick reference place to return to for the key definitions.

8.  Use of "We..." was especially confusing when describing the steps.  For example, line 147 states "We then remove one node from..."  The user doesn't actually do this; it is automated.  The current text makes it seem that there are many many analysis steps that have to be done manually by the user.

Thanks. We have changed the terminology to differentiate when the tool automates steps versus when the user has to do something. For example, the end of the methods sections describes the three inputs the user provides.

9.  It was very hard to read the text in Figure 3, which could be problematic since the authors refer extensively to node and link names later in the paper.  Is there a "vector" version of this graphic, where the text can be clearly read at any resolution?

Figure 3 has been recreated so that node labels are easier to read. The figure also uses different symbols (rectangles, triangles, dashed triangles, etc.) to represent service areas, reservoirs, proposed reservoirs, and other nodes. The arrowheads have also been enlarged.

10.  When presenting the result, I found lines 306-310 to be redundant; the authors are simply restating the steps of the method.  Instead I would like to see a clear discussion of how to actually interpret the PCP plot in figure 4.  What do the colors mean?  What is the difference between an extracted centrality of 0, and a low extracted centrality of less than 5?  Right now, figure 4 has a very short caption and a lot of colored lines so it will be hard for a reader to find any meaningful interpretation without looking at the text.

Thanks. We have added more discussion to this paragraph to describe the colors (type of node: reservoir, service area, junction, sink, and source), extracted centrality range, and how to interpret the parallel coordinate plot to reach general, visual, and qualitative observations regarding the performance of various classes of nodes.

11.  The colors are almost impossible to discern in Figure 5.  This is especially true between "3rd most" and "3rd least".  Perhaps four colors could be used, one color for the "most" stable, another for "most" topologically significant, a 3rd for "least" stable, and a 4th for "least" topologically significant.

We now use only four colors in the revised Figure 5: most and least for stability and topological significance.

12.  [line 362-363] Are there some literature citations about managers' opinions on redundancy?  From my experience redundancy greatly improves the security of a network, so it's a desired trait.

Correct, redundancy can also be seen as a positive characteristic of a network, particularly that it offers water managers operational flexibility. We have revised this section and now discuss how quantifying redundancy can help identify areas of operational flexibility as well as areas where it may be possible to remove nodes and save money and other resources. The text no longer makes an implicit or explicit judgment about whether redundancy is good or bad. We could not find a citation for managers’ opinions on redundancy. Do you have suggested citations? This is an example of why quantification represents a key contribution – because we didn’t find any other data or literature quantifying system redundancy.

13.  The citation to Matrosov et al. (2011) is included in the references but not mentioned in the text.

Thank you; this citation has been removed.

Reviewer #3: Review of "Identifying stability, topological significance, and redundancies in water resource networks using parallel coordinate plotting"  
  
The manuscript analyzes the effects of changes to water resources network structure and management on (1) node stability, (2) topological significance, and (3) node redundancy using parallel coordinate plots. The method is tested on a 56-node network in Idaho/Utah. Nodes with low topological significance may be good candidate locations to remove dams, transfer water, or implement conservation measures.  
  
The manuscript as it stands is not suitable for publication, for the reasons detailed below. The authors have already anticipated many of these concerns in the "Future Work" section. Once completed, these additional analyses will greatly improve the ability to interpret the results for water resources management applications.

The main contribution of this method and tool is to identify key network nodes through quantification and ranking of stability, topological significance, and redundancies that then should be used to focus detailed simulation modeling. The revision now includes flow direction. We agree with the reviewer that including flow magnitude would provide additional information but also believe that the method as presented with flow direction provides significant policy relevant information and significantly narrows options for managers to explore with detailed simulation modeling. The revised manuscript now emphasizes the main contributions of quantification of the performance metrics and interpretation to apply the tool to water resources network management instead of the parallel coordinate plots which were already produced by Singer and Greenshpan (2009). We also expanded discussion the section to explain more ways the method and tool can be utilized in policy making (e.g., environmental protection, conservation measures, and monitoring).

Comments:  
  
1. The core of the analysis, computing the stability and topological significance of network nodes, appears trivial. According to the given definitions, the most unstable nodes will have degree 1, and the most toplogically significant nodes will have the maximum number of inbound connections (including also the inbound connections to adjacent nodes). These are common problems in graph theory applications, and writing a program to obtain these results is not a novel contribution unless significant water management insights can also be obtained.

Yes, the most unstable nodes in the network are the nodes with degree 1 which may be intuitive. However, the ranking method and tool presented go way beyond this trivial observation for a simple case and identify, rank, and compare nodes that are neither the most nor least stable. The Bear River network example illustrates the importance of this extension. Yes, the most stable nodes are sources (i.e., they don’t depend on any other node for water) but many nodes in the network have two or multiple sources. For example, the Weber and Cache Valley Urban service areas each have two sources (a reservoir and an inflow), but Weber is the 52/55 for stability and while Cache Valley Urban ranks 25/55.

Topological significance is not synonymous with the maximum number of connections. The most topologically significant nodes are those that when extracted create the most instability in other nodes. Topological significance nodes can be those with the most connections but that is not required. In the Bear River example, the Corinne Junction has five upstream connections, two downstream links, and a topological significance rank of 11th compared to Junction 32-88 that has a topological significance rank #1 but only 4 upstream and 1 downstream connections. Similarly, Junction 48-49 has a topological significance rank #6 but only 1 upstream and 3 downstream connections. We now include text in the results section that describes these findings and points out how topological significance is not synonymous with the maximum number of connections.

Topological significance also offers significant water management insights compared to simply counting the number of connections to a node. Topological significance is a measure of how a particular node influences the other nodes in the water network. This influence is seen when extracting the node from a network and a factor of both the number of connections as well as the nature of the connections (i.e., being the only connection for a branch of the system) factor into TS. In the Bear River example, J32-88 is the most significant node; geographically, it connects the Weber and Wasatch branch to the rest of the network. Hyrum Reservoir is ranked 6th because it is located upstream of most of the network though it only has two junction connections and evaporation.

2. Significant water management insights cannot be obtained from the analysis in its current form because it does not consider the volume and direction of flows. For example, a node with low topological significance is not necessarily a good candidate for water transfers or dam removal, etc., unless its volumetric contributions to water supply are also considered. Connectivity alone is not a sufficient measure to derive management insights. The authors have anticipated this problem in the "Future Work" section of the manuscript and have provided a detailed plan of how to incorporate volume-weighted measures of node centrality. This additional analysis will allow water management insights to be derived.

Yes, agreed, and our earlier response to the initial overview comment also applies here. To recap, the analysis now considers flow-direction and we have expanded the discussion section to describe how the method and tool can be used considering flow-direction to advance water resources network management.

3. The manuscript seems to overemphasize the value of parallel coordinate plots (PCPs) to perform this network analysis. These plots do provide useful visualizations of network topology. However, the key results in the manuscript (i.e., a ranking of nodes according to topological significance) can be obtained just as easily without PCPs.

Thank you for pointing out that the manuscript incorrectly emphasizes parallel coordinate plots. We have revised to emphasize the main contribution of this work as quantitative ranking each network node. The ranking still requires the centrality values used to construct the parallel coordinate plots, but not the plot itself. The plot is principally used to visually illustrate and define key attributes of node stability, topological significance, and redundancy (e.g., drops in traces, locations where traces drop, similar traces, etc.). We retain the parallel coordinate plot as we believe the plot helps the user better conceptually understand the performance metrics. However, the abstract, introduction, and section 2 now more clearly describe the plot as part of prior work and that the contribution from our tool is to rank nodes and apply this ranking to inform system management.

4. The method relies on calculating the centrality of nodes before and after the extraction of other nodes. What about extraction of multiple nodes? For example, the sequential extraction of one node followed by another may cause much different network effects than removing either node individually.

Yes, the simultaneous extraction of multiple nodes could offer interesting results. There are several significant issues to resolve to move forward with multiple-node extractions. These issues include which nodes to group together, how to quantify topological significance (calculate a group topological significance?), how to identify individual node topological significance from the group ranking, similarly for stability, how to rank group and individual nodes, and how to identify redundant node pairs/groups. At the same time, results from the single-node extractions presented in the paper can give insight into multiple-node extraction: (a) removing multiple nodes with high topological significance would cause more instability in the network than removing a node group with less topological significance, (b) removing a group of nodes that are start nodes for unstable nodes will create even greater instability, and (c) removal of a group including any nodes identified as a redundant node pair will reduce redundancy.

5. In general the resolution of the figures should be improved.

Figure 3 and Figure 5 have been changed and are now submitted in vector form. The parallel coordinate plots in Figure 2 have also been modified.

6. Figure 5, the plot showing the results superimposed on the network map, is a good idea and is very helpful.

Thank you.

Reviewer #4: This is a great paper! NEVIS a sophisticated method for Network Visualization based on ||-coords is used and extended very  
capably to solve a difficult problem in a real water resources network. The results are important and impressive.  
  
Minor but important changes  
1. Separate entry for the 2009 book on ||-coords (In addition to or instead of the 1985 reference)

Each reference (2009 or 1985) has been used as appropriate.

2. Unlike Pie Charts or Bar Charts etc ||-ccords are a Multidimensional COORDINATE SYSTEM and not just a "plot" i.e. PCP  
The modern notation is ||-coords.

The text has been changed so that parallel coordinates is used when discussing the coordinate system and parallel coordinate plot when talking about the plots themselves. We decided not to introduce any further annotation or acronyms for the reader to remember throughout reading the manuscript, so the words “parallel coordinates” or “parallel coordinate plot” are now used throughout.