Calculating a Distance Matrix of Locations Along River Networks

These instructions enable you to measure the length of line segments between points, which is much more useful than simple straight-line distances when you need to work in a network environment such as along water courses, as well as the more traditional applications on roads and utility features. Using ESRI's ArcInfo, ArcView 3.x and Microsoft's Access software, perform the following steps to obtain a pair-wise distance matrix between locations :

- Convert line shapefile to coverage and clean/build line topology
- Calculate distances between origin and destination sites along the river network
- Create a cross tabulation query to organize the data into a matrix

You need to install the extension **Fastest Shortest Network Paths v1.0** created by Soeren Alsleben (available for free at <u>http://arcscripts.esri.com/</u>) – this requires ESRI's **Network Analyst** to also be installed (available with paid license) – in ArcView 3.x.

ORIGINAL DATA

Rivers.shp	line shapefile of river branches and tributaries
Sites.shp	point shapefile of sampling/station locations with
	unique identification field values for each record

CREATED DATA

Rivers	coverage converted from the shapefile
SNP_sites.shp	shapefile consisting of pair wise distance segments
	between selected site locations along rivers
SNP_distances.txt	exported table of the results

Steps in ArcInfo Workstation:

Topology explicitly defines spatial relationships. The major topological concept associated with linear features, such as rivers, is connectivity. Arc-node topology defines connectivity: arcs (linear features) are connected to each other if they share a common node (endpoints and intersections). This is the basis for many network tracing and path finding operations.

- 1. Start Arc/Info
- 2. At the command prompt, type **arctools**
- 3. Select COMMAND TOOLS and click OK
- 4. Choose FILE >>> MANAGE WORKSPACES
- 5. Navigate to your working subdirectory (e.g. c:\workspace\river_distances) and click OK
- 6. Click on the CONVERSION pull-down menu
- 7. Choose TO ARC >>> SHAPE TO ARC

Мар	Tools		٦
Edit	Tools		
Grid	Tools		_
Comm	and Too	ols	
Land	Record	ls	
4			

8. Right click in the "Input ArcView shape:" text box

Input ArcView shape:	c:\workspace\river_distances\rivers	
Output coverage:	c:\workspace\river_distances\rivers	
	Subclass:	

- 9. Select the **Rivers.shp** file as the input
- 10. Double click to highlight the input text box and simultaneously hit the **Ctrl-C** keys on the keyboard to copy the information
- 11. Click in the "Output coverage:" text box and simultaneously hit the Ctrl-V keys on the keyboard to paste the information
- 12. Click OK
- 13. Click on the EDIT pull-down menu
- 14. Choose TOPOLOGY >>> BUILD FEATURES
- 15. Right click in the "Coverage:" text box and select the **Rivers** coverage as the input
- 16. Click on the CLEAN option button
- 17. Click a check in the "Clean options:" check box to optionally change the defaults
- 18. Highlight the LINE feature class
- 19. Click APPLY



Coverage Topology	<u>_ </u>
Coverage: c:\workspace\river_dis	tances\rivers
P	
LINE Subc	lass:
C Build @ Clean C Idedit	
, pulle, clean, lucut	
Clean options:	
	river distances\rivers
Output coverage: c:\workspace\	
Output coverage: c:\workspace\	
Output coverage: c:\workspace\; Dangle length: 0	
Output coverage: c:\workspace\; Dangle length: 0 Fuzzy tolerance: 15042.2328	
Output coverage: c:\workspace\; Dangle length: 0 Fuzzy tolerance: 15042.2328	

- 20. Click on the BUILD option button
- 21. Highlight the LINE feature class
- 22. Click APPLY
- 23. Click CANCEL when finished
- 24. Choose ARCTOOLS >>> QUIT
- 25. QUIT from the ArcTools interface and then type '**q**' to quit ArcInfo altogether

You now have a line **coverage** that is clean (corrected for geometric coordinate errors by splitting lines at intersections, etc.) and built (created/updated the feature attribute table based on the changes made by the clean operation) with arc-node topology for the rivers data – very important for the network analysis below.

Steps in ArcView 3.2:

FAST SNP V1.0 for ArcView 3.x (Network Analyst required) is a third-party extension that calculates shortest paths between origins and destinations using the path length or travel time.

- 1. Start a new project in ArcView 3.2
- 2. Open a new View and add the **Rivers** coverage and the **Sites.shp** theme
- 3. Choose FILE >>> EXTENSIONS
- 4. Click a check beside the Network Analyst and Fastest Shortest Paths v1.0 extensions
- 5. Turn the themes on
- 6. Highlight the river theme 7. Click on the Fast SNP v1.0 button 8. Select **Sites.shp** as the point theme Q Fast SNP v1.0 x representing origins and destinations Please select the point theme representing origins and OK destinations. 9. Click OK 10. Select the Point-ID field and click OK Sites.shp Cancel 11. Select the Point-Label field and click OK 12. Select the origin(s); for a 🔍 Fast SNP v1.0 X distance matrix of all possible Please select the origin(s) OK pair-wise combinations. choose EACH RECORD *** EACH RECORD * Cancel 13. Click OK 1 (ID: 1) 10 (ID: 10) A Fast SNP v1.0 × 14. Select the 11 (ID: 11) destination(s); Please select the destination(s). 12 (ID: 12) OK again, use EACH 13 (ID: 13) RECORD for all *** EACH RECORD * Cancel 14 (ID: 14) 1 (ID: 1) 15 (ID: 15) possible pair 10 (ID: 10) distances 11 0D. 111 Q Fast SNP v1.0 15. Select the × network theme Please select the Network-Theme OK (e.g., **rivers**) and click OK Cancel Rivers 16.Enter Fast SNP v1.0 X а Please enter the search tolerance (points <-> network). 0K Select 'Cancel' for the default value (currently [1.50422e+006]). Γ Cancel

search tolerance or click CANCEL for the default

- 17. Click OK to the warning that the distance will be used as the cost for calculating shortest paths
- 18. Enter the directory and filename (e.g. c:\workspace\river_distances \SNP_sites.shp) for the resulting shapefile and click OK
- 19. Click YES or NO to exclude 'A-A' paths (identical origin/destination locations)







The Origin and Destination fields indicate the site pairs and the Cost field shows the distance in map units along the river and between the sites. 22.Choose FILE >>> EXPORT

- 23. Select DELIMITED TEXT as the export format and click OK
- 24. Save as c:\workspace\river_distances\SNP_distances.txt

Steps in MS Access:

The shapefile that results from the network analysis simply lists the pair-wise distances along the river network. The following assumes prior familiarity with starting an Access database and instructs on how to get the data into a matrix form using this common database software program:



1. Start MS Access by creating a new Blank Access database; e.g.

c:\workspace\river_distances\SNP_distances.mdb

Create a new table by importing the SNP_distances.txt file.

- 2. Choose INSERT >>> TABLE
- 3. Select IMPORT TABLE and click OK

Import	<u>? ×</u>
Look in: 🚺 river_distances	💽 🖕 🔁 🔕 🗙 📸 🏢 - Tools -
History History My Documents	
Desktop Favorites	
My Network Places File name: Files of type: Text Files	✓ Import Cancel
4. Navigate to your working direc	tory
5. Select "Files of type:" IEXI FIL	_ES
Your data seems to be in a 'Delimited' format. If it isn't, choose to describes your data.	he format that more correctly Tifield ween each field Ween each field NEXT Note led to be led t SNP_distances.txt 7. Click IMPORT 8. In the Import Text Wizard, accept DELIMITED and click NEXT
Sample data from file: C:\WORKSPACE\RIVER_DISTANCES\SNP_ 1 "Origin ID", "Destination ID", "Origin	DISTANCES.TXT.
2 53,54,53,54,22729914.000 3 53,52,53,52,23744546.000 4 53,51,53,51,83578992.000 5 53,50,53,50,98885544.000 6 53,49,53,49,104177888.000 7 53 48 53 48 117182608 000	What delimiter separates your fields? Select the appropriate delimiter and see how your text is affected in the preview below. Choose the delimiter that separates your fields: C Tab C Semicolon C Tab C Semicolon
	First Row Contains Field Names Text Qualifier:
Advanced Cancel <back< td=""><td>Origin ID Destination ID Origin Label Destination I 53 54 53 54 4 53 54 53 54 4</td></back<>	Origin ID Destination ID Origin Label Destination I 53 54 53 54 4 53 54 53 54 4
delimiter, click a check by "First Row Contains Field Names," and select the double quote (") as the "Text	53 52 53 52 53 51 53 51 53 50 53 50 53 49 53 49 53 48 53 48 53 3 53 6
<i>Qualifier</i> ", and then click NEXT	Advanced Cancel < Back Next > Finish

10. Click NEXT subsequently to accept the defaults (or optionally modify) and then click FINISH when done

11. Open the new table to view it



It is virtually identical to the attribute table from the SNP sites.shp file. Apply a query to create a cross tabulated matrix. A crosstab query displays values (and optionally sums, counts. and averages) from one field in a table and groups them by one set of record IDs listed down the left side of the datasheet and another set of record IDs listed across the top of the datasheet.

12. In the Database window, click QUERIES under Objects.

- 13. Click NEW on the Database window toolbar
- 14. In the New Query dialog box, click CROSSTAB QUERY WIZARD
- 15. Click OK
- 16. Follow the directions in the wizard dialog boxes and click NEXT after each selection: Crosstab Query Wizard
 - Table = SNP_Distances
 - Row headings = Origin ID
 - Column headings
 = Destination ID
 - Specify the field for each column-row intersection = Cost

The default Avg function is okay since the table contains the same Cost for A-B and B-A pairs. 17. Click FINISH 18. View the query This final display of the

What number do you want calculated for each column and row intersection?		ields:			Functions:
each column and row intersection? For example, you could calculate the of the field Order Amount for each employee (column) by country and ro (row). Do you want to summarize each row? Ves, include row sums.	sum gion	D Drigin L Destina	abel Ition Label		Avg Count First Last Max Min StDev Sum Var
Sample:					
	Origin ID		Destination I	Destination	I Destination I
	Origin ID1	ł	Avg(Cost)	Q.	
	Origin ID2				
	Origin ID3				
	Origin ID4	Î			
		1		r	

data in matrix format indicates Site IDs in the column and row headings, with the distance between them along the river network in the corresponding table cell.