**Water Data in Time and Space**

Prepared by David R. Maidment

Waterways Centre for Freshwater Research

University of Canterbury

Based in part on material from Brent Watson

Horizons Regional Council

Palmerston North

Revised with corrections, 8 March 2018

March 2018

**Goals of the Exercise**

This exercise shows how hydrologic and water quality data are obtained using water data services, and how to create and work with a stream network for the Selwyn catchment.

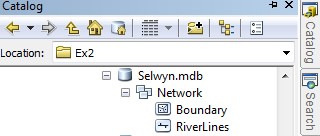
**Computer and Data Requirements**

To carry out this exercise, you need to have a computer, which runs ArcGIS Desktop version

10.5. This exercise will also work with version 10.4.1 if you do not have access to Version

10.5.

In the first part of this exercise using ArcGIS Desktop, you will be working with the following Geodatabase called **Selwyn**, which has a Feature Dataset within it called **Network**, and within that there are two feature classes **Boundary** and **RiverLines**, which are the Boundary of the Selwyn catchment and the River Environment Classification RiverLines for that area.



You can get these data from this zip file:

<http://www.caee.utexas.edu/prof/maidment/Canterbury/Ex3/Ex3Data.zip>or from the UC Learn web site. You need to establish a working folder to do the exercise. This can be in any convenient location on the computer you are working on (e.g. …\Ex3\Ex3Data). After you have downloaded the zip file **Ex3Data.zip** double click on the file and you should see Winzip or other zip utility to open the file on your computer (if it doesn’t open you’ll have to unzip this file on a computer that has a zip utility installed).

# Part One: Water Data in Time

When you are querying the LAWA web site for data or charts, what is actually happening behind the scenes is that queries are being made using a protocol called the Sensor Observations Service and the results are being returned in a language called WaterML2. These are standards of the Open Geospatial Consortium, whose protocols have been adopted by the Regional Councils and Crown Research Institutes in New Zealand to support open data sharing. We are hoping that the data services from these organizations will be opened up to all users but in the meantime, the following examples of water data services access from the Horizons Regional Council for data from the Mangatainoka River at Pahiatua Town Bridge are offered as examples.

## Getting Flow Data from Mangatainoka River at Pahiatua Town Bridge

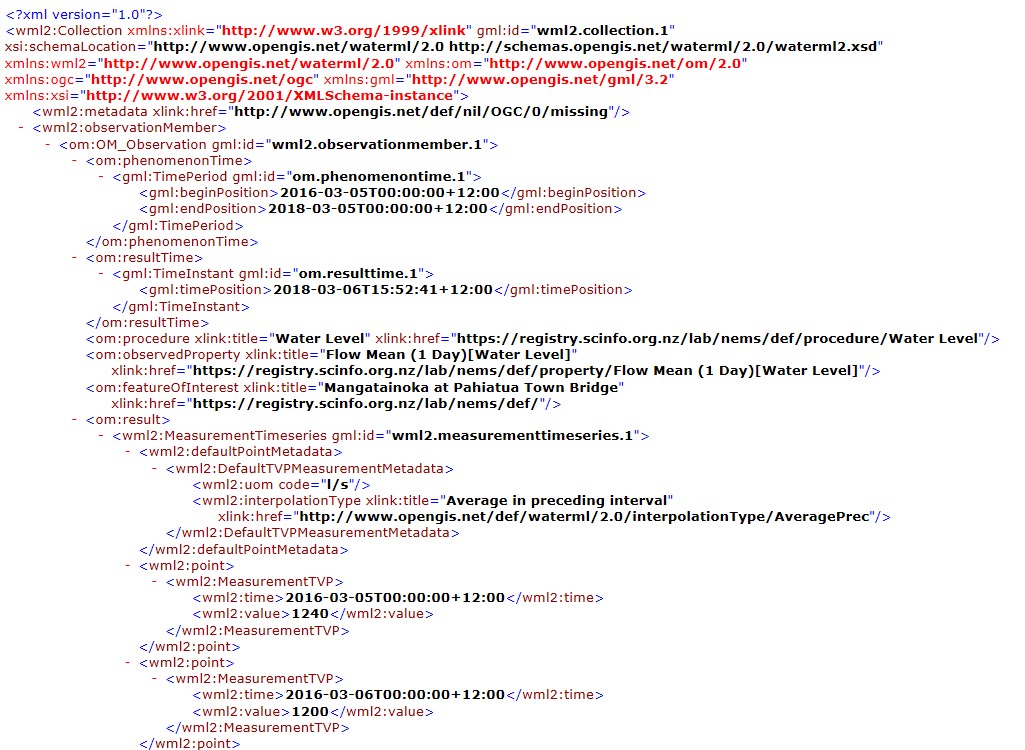
This example enables automated acquisition of daily mean flow or discharge data for the past two years. You will make a “Get Observation” request of the Sensor Observation Service (SOS) operated by the Horizons Regional Council, and you are using a service that has 4 parts:

* **Feature of Interest:** Mangatainoka at Pahiatua Town Bridge
* **Observed Property:** Flow Mean (1 Day) – this means the daily mean value of 5 minute observations
* **Procedure of Measurement:** [Water Level] – measurements of water level that have been converted to flow using a rating curve
* **Result:** Temporal Filter of P2Y which means the past 2 years backwards from the time that the request is made

Copy the text below, and launch the following web query from a web browser:

http://tsdata.horizons.govt.nz/contactrec.hts?Service=SOS&Request=GetObservation&FeatureOfInterest=Mangatainoka at Pahiatua Town Bridge&ObservedProperty=Flow Mean (1 Day)[Water Level]&TemporalFilter=om:phenomenonTime,P2Y

The result will emerge like this:



*To be turned in: A screen capture of the header of your WaterML response for flow along with the first couple of data values, as shown above. What is the time period of your data request (from date, to date)? What are the units of the flow data? What was the mean daily flow of the Mangatainoka River at Pahiatua Town Bridge on 5 March 2018?*

## Getting E. Coli Data from Mangatainoka River at Pahiatua Town Bridge

In a similar way as for the flow data, launch the query:

http://hilltopserver.horizons.govt.nz/cr\_provisional.hts?Service=SOS&Request=GetObservation&FeatureOfInterest=Mangatainoka at Pahiatua Town Bridge&ObservedProperty=E. coli by MPN (HRC)[E. coli by MPN (HRC)]&TemporalFilter=om:phenomenonTime,P2Y

and the result appears as:



*To be turned in: A screen capture of the header of your WaterML response for E. Coli along with the first couple of data values, as shown above. What is the Feature of Interest, Observed Property, Procedure of Measurement and Temporal Domain of the Result? What is the time period of your data request (from date, to date)? What are the units of the E. Coli data?*

## Using Water Data Services from Excel

As you would appreciate, web services are great for computers but not easy to interpret for hand computation. Excel knows how to read XML documents and to parse the information they contain into data fields.

Open Excel and w

ithin a blank worksheet, select the

**Data**

ribbon and

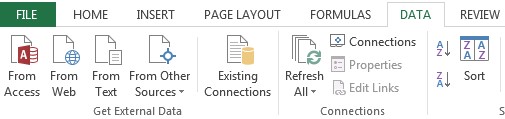
**From Web**

function

within the

**Get external Data**

toolset.



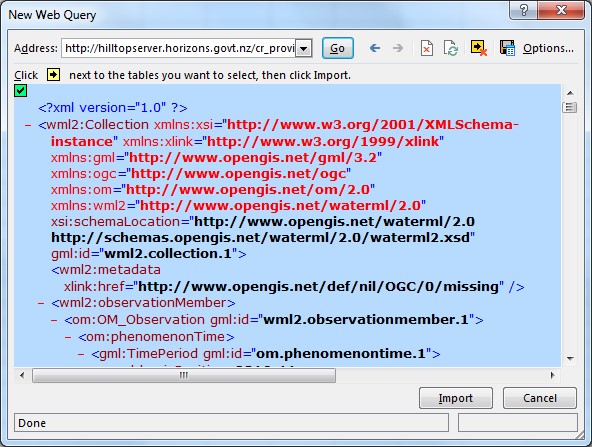
A popup web-query will default to your default webpage.



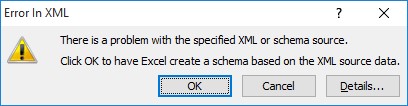
Copy and paste the same web query that we used for E Coli previously into the address bar where [http://www.canterbury.ac.nz](http://www.canterbury.ac.nz/) appears above. Here is the web query again:

http://hilltopserver.horizons.govt.nz/cr\_provisional.hts?Service=SOS&Request=GetObservation&FeatureOfInterest=Mangatainoka at Pahiatua Town Bridge&ObservedProperty=E. coli by MPN (HRC)[E. coli by MPN (HRC)]&TemporalFilter=om:phenomenonTime,P2Y

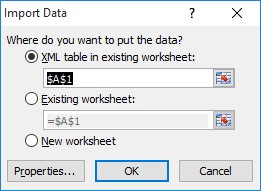
Hit **Go** and you’ll see the same XML response that you saw before come up. Click on the little yellow arrow in the top left corner of display area so that it turns into a green check mark, as shown below, and hit **Import**.



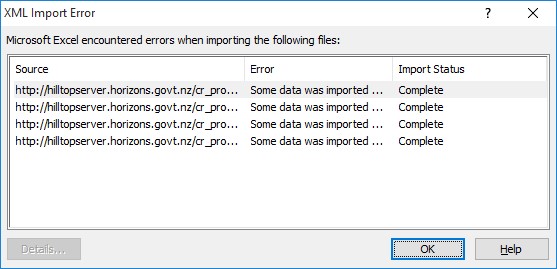
Excel doesn’t know how to interpret all the header information in the WaterML2.0 Schema; however we are happy to allowing Excel to create a new schema from the (time, value) pairs of data in the repeated rows below the header. Select **OK** when you see the query below.



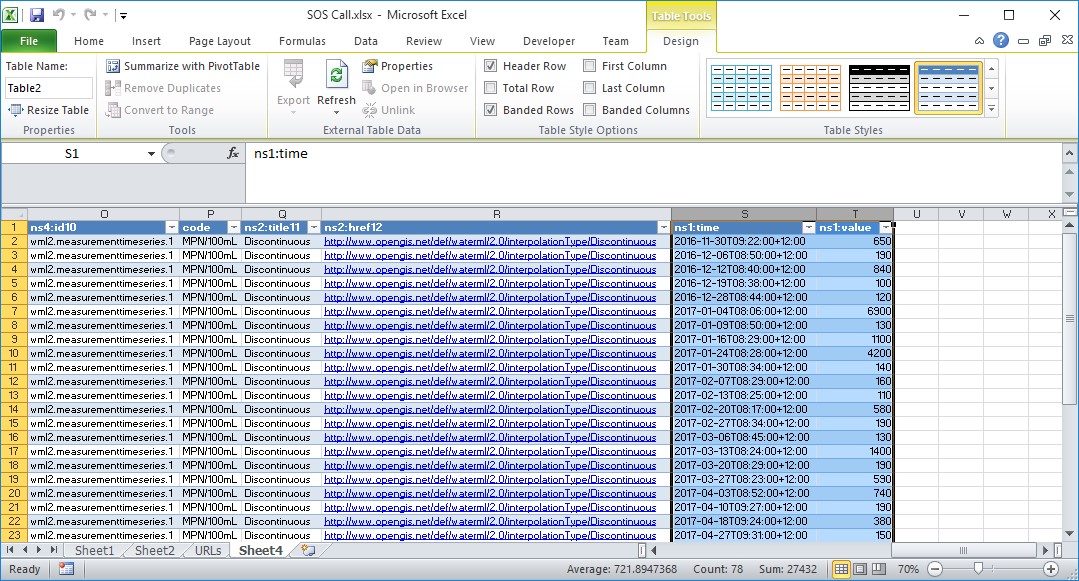
Select where in the workbook you want to place the data.



Excel will report at the importing steps and display any errors.



The final import contains a number of columns due to the complexity of the WaterML2.0 schema. Column S & T contains the (Time,Value) pairs.

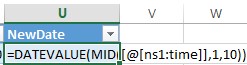


note: the date &time is provided in Universal Time Coordinate system format.

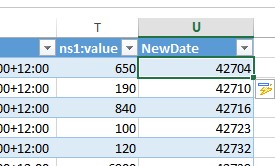
Now you have time series data. The time stamp in column S can be converted into a normal Excel date as follows. Create a new column called **NewDate** and set the resulting formula in the first row of this new column as:

=DATEVALUE(MID([@[ns1:time]],1,10))

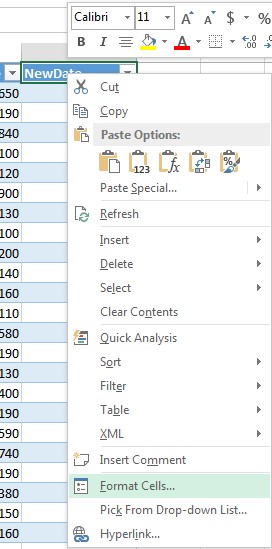
This query is using two Excel functions, DATEVALUE and MID, one inside the other. The MID function selects the text value from the column that it is pointing to (the time field), and isolates the first 10 characters in that field, 2016-11-30, and then the DATEVALUE function converts these into an Excel Date format.



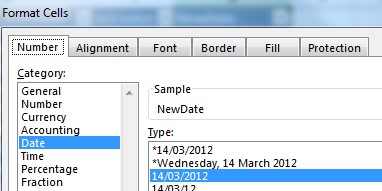
You’ll see that Excel creates a column of integers that look incomprehensible. However, if you divide the first number, 42704 by 365.25 you’ll get the result 116.9172, which means that this value is 116.91 years from January 1, 1900. This date is actually November 30, 2016 so you can see the connection.

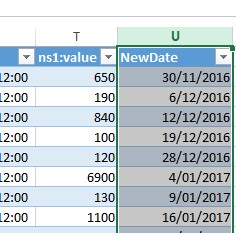


Right click on this column and select **Format Cells**

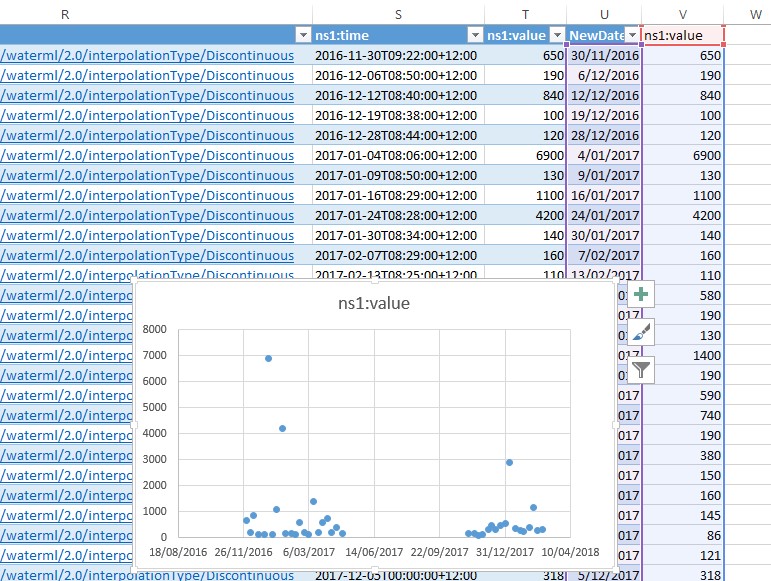


And you’ll get a nicely formatted date field.

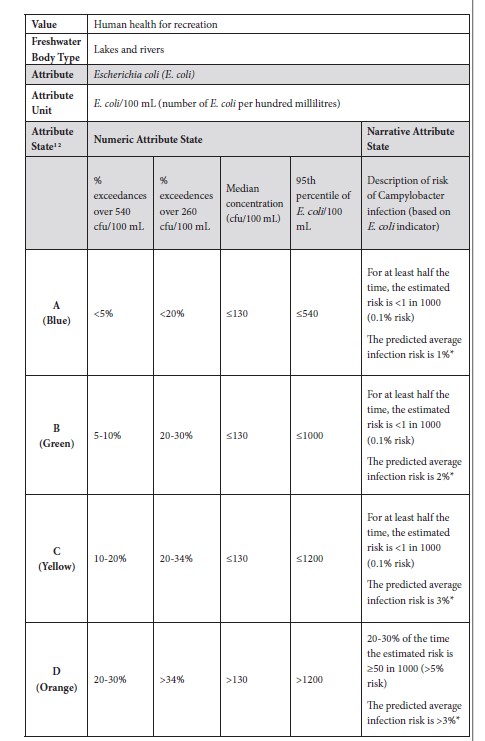




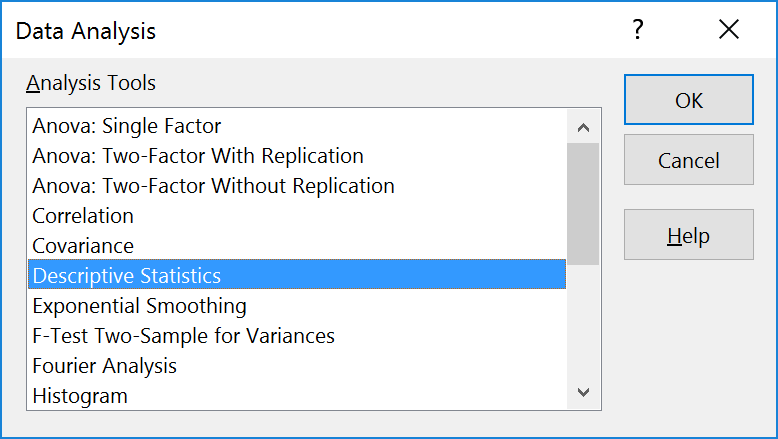
Plot a nice time series chart of these data and find the critical percentile values of these data that correspond to the New Zealand coliform data standards.



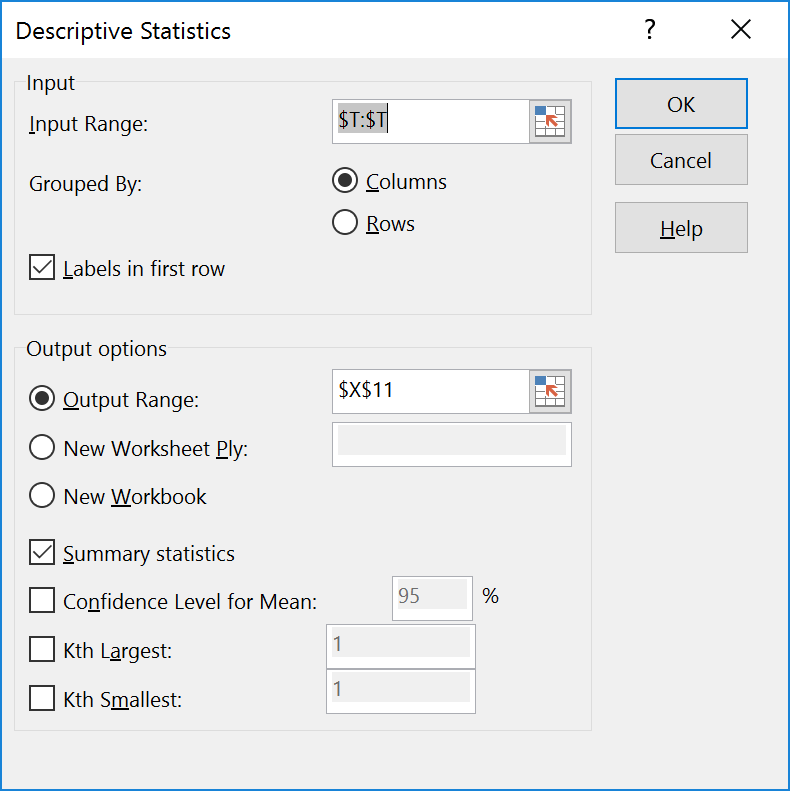
Below are the New Zealand human health standards for E. Coli. These are taken from p.39 of the **National Policy Statement for Freshwater Management, 2014**, which is posted on the course web page along with this assignment. Let’s see how the data we’ve downloaded conform to these standards.



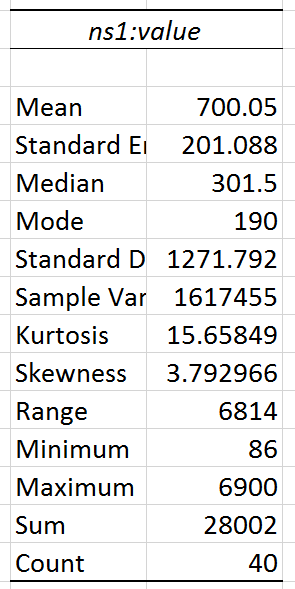
First, do some descriptive statistics of these data using the Excel Analysis Toolpak. Choose the descriptive statistics option



Ask for Summary Statistics

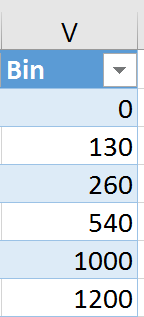


And you’ll get a result like this:

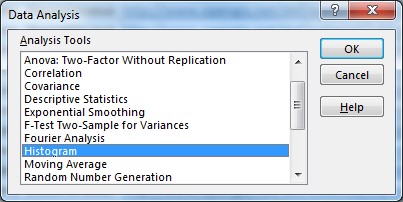


From which you can see that the median E. Coli concentration of the 40 values downloaded over the past two years is 301 CFU/100mL.

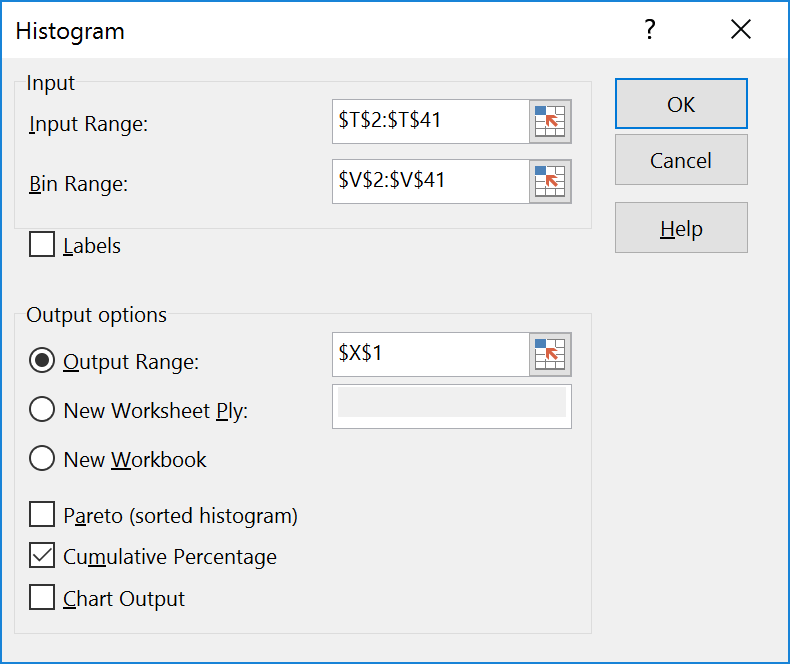
Next, we’ll construct a histogram of these data by using the Excel Analysis Toolpak option Histogram. Before we do that, we need to create a Bin range of Coliform concentrations that are used in the New Zealand water quality standards for E Coli:



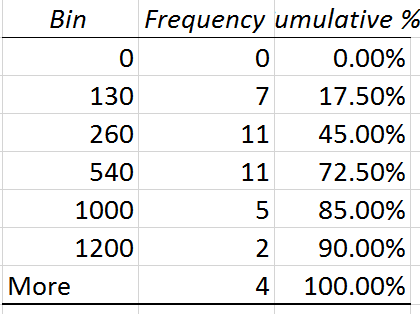
Then open the Histogram function:



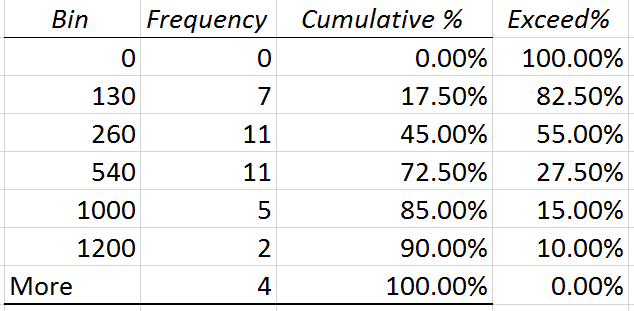
Using the following options (click on Cumulative Percentage but don’t click on Labels)



And you will get a cumulative frequency curve of the E. Coli data:



You can get the % of the data that exceed these values by taking 1 – Cumulative %



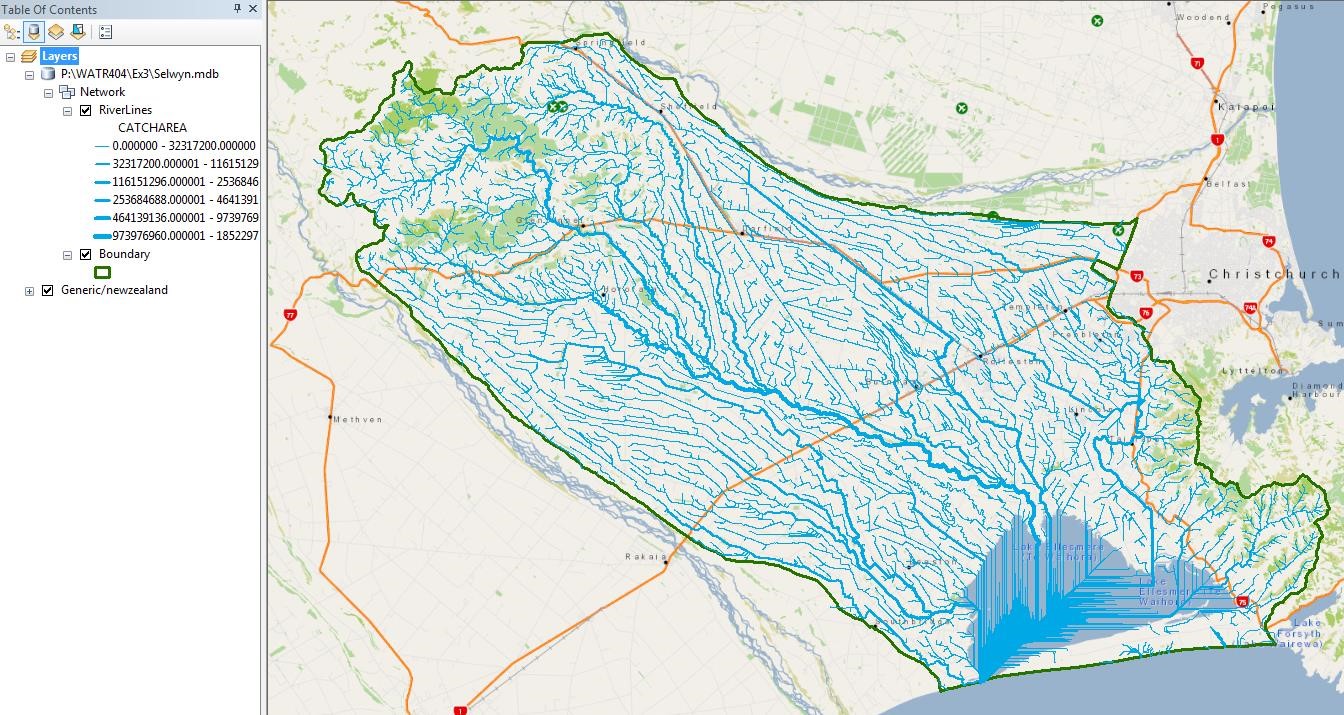
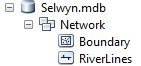
From this, you can see that about 83% of the data exceed 130 CFU/100mL, 55% exceed 260, 28% exceed 540, 15% exceed 1000, and 10% exceed 1200. Oops! Doesn’t look too good for water quality at this location.

*To be turned in: Plot a nice time series chart of these data and find the critical percentile values of these data that correspond to the New Zealand coliform data standards. What quality level do these data represent?*

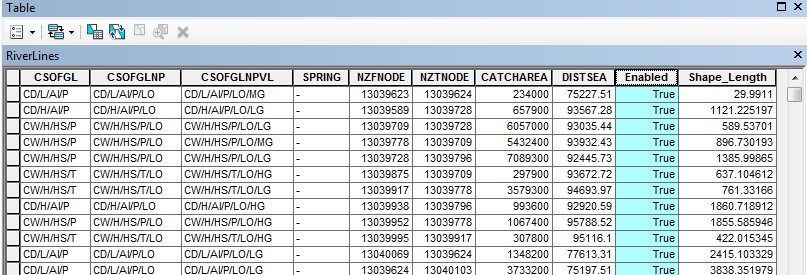
# Part 2. Water Data in Space

## River Network for the Selwyn Catchment

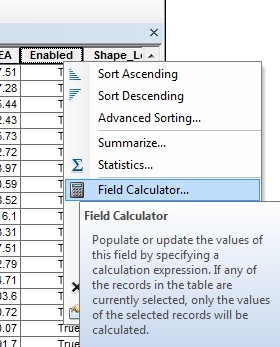
Open ArcMap, select a base map, open the data file **Ex3Data.zip** and from the **Selwyn** geodatabase, add the **Boundary** and **RiverLines** feature classes to the map display. Symbolize the river lines using the **CatchArea** attribute.



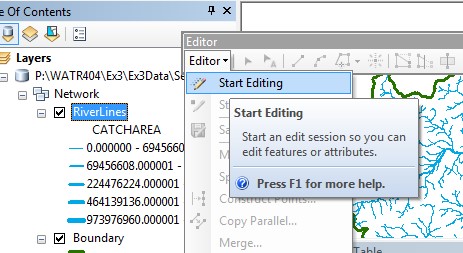
Open the Attribute Table for RiverLines and check if the **Enabled** field is set to **True**. If so, the geometric network cannot be formed with this feature class.



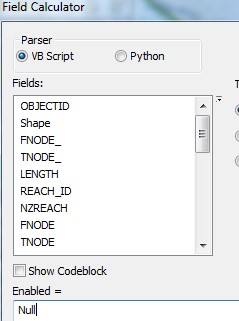
To avoid this problem, do a calculation that sets the Enabled field to **Null** and then you’ll be good to go with these or other river lines that you want to make into a network.



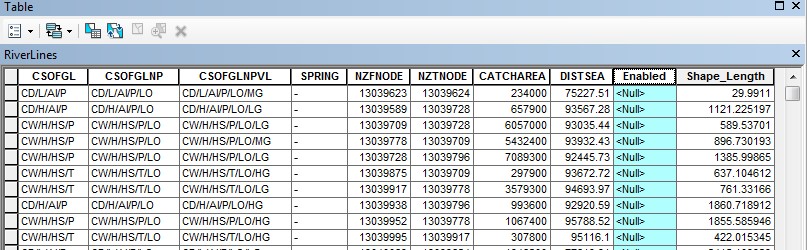
If the field calculator is greyed out, right click of on the right of the ArcMap ribbon, turn on the Editor tool and start editing the data



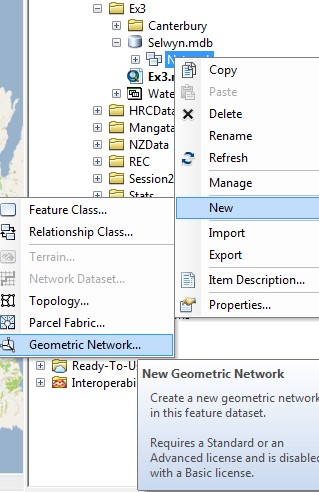
Here is how to set the calculation in the field calculator:



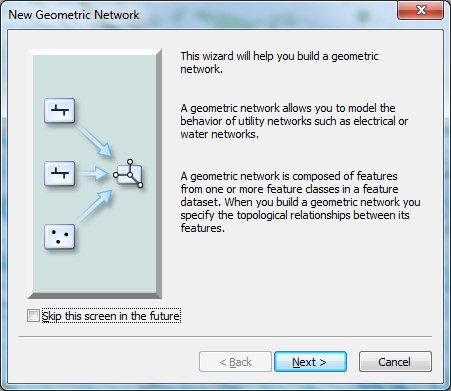
and you get this result, as needed.



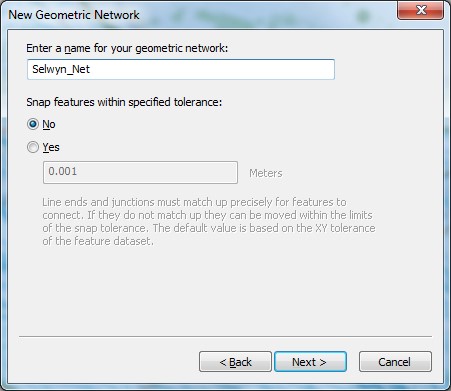
In ArcCatalog, Right click on the **Network Feature Dataset** and create a **New Geometric Network**



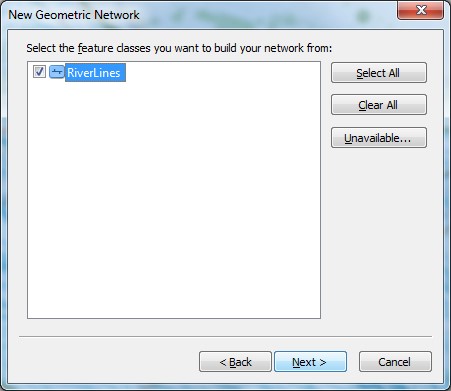
You’ll see the following screen pop up, and hit **Next**



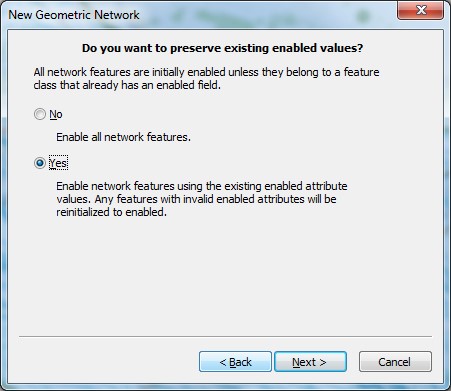
Name your network **Selwyn\_Net** and hit **Next**



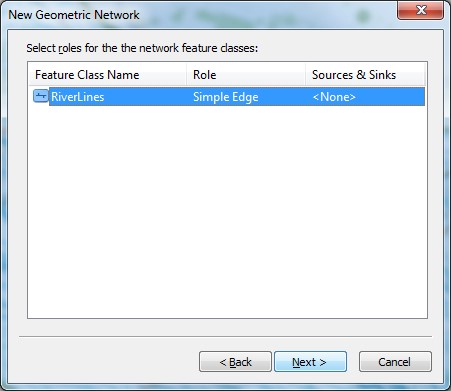
Select **RiverLines** to participate in the network and hit **Next**



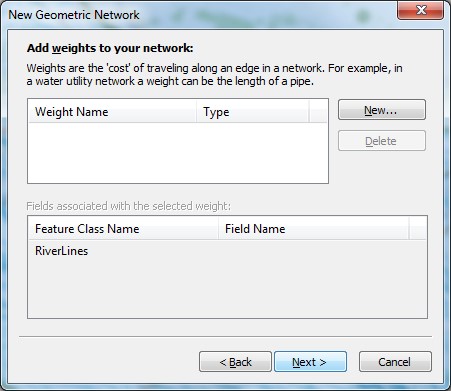
and hit **Next** again



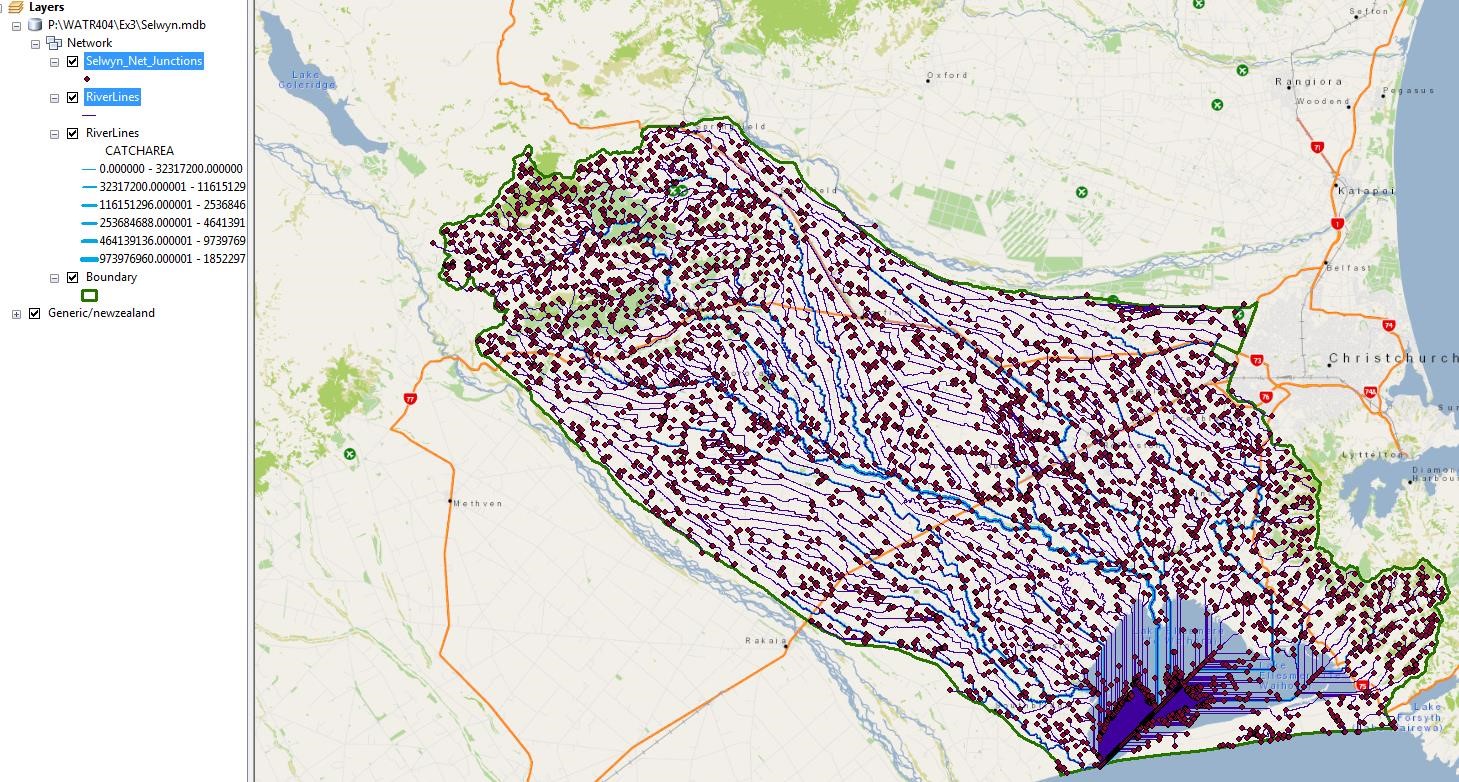
Select **RiverLines** to be your **SimpleEdge** feature class in the network



Hit **Next**



and lastly **Finish.**  You’ll see the computer think for a while and then come up with a new Geometric Network that has Junctions to connect every Edge. What has happened is that each RiverLine feature now knows what river lines it is connected to.

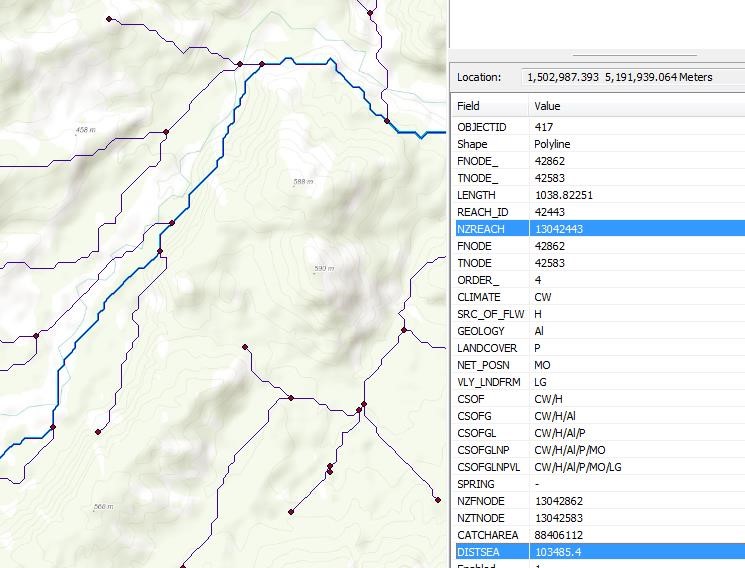
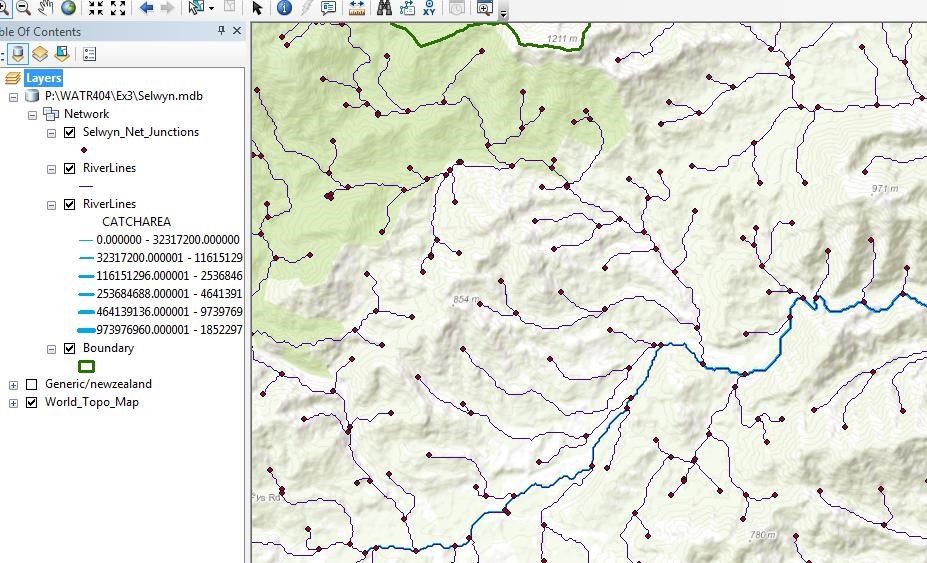


There seems to be a systematic displacement between the REC river lines and the river line images below in the **Generic New Zealand** base map we’ve been using, so let’s use the **World Topo Map** basemap instead and zoom in to a particular area within the network.

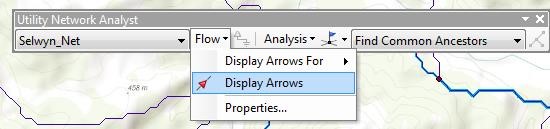
tool, you’ll see a set of attributes for each reach that includes its **NZREACH** number that is unique for all of the 600,000 reaches in New Zealand, and also **DISTSEA** which is the distance downstream from this reach that the water flow reaches the sea in meters. In the example shown this is 103485 meters or 103.485 km.

If we do a query us

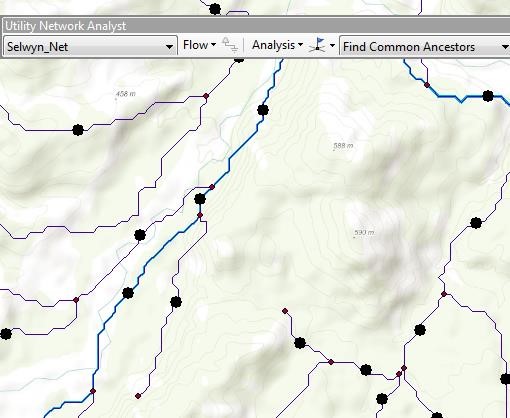
ing the



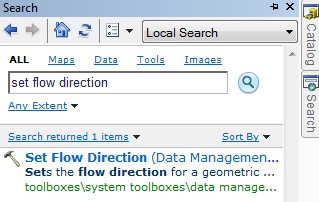
Right click on the grey area on the right hand side of the top ribbon in ArcMap and select **Utility Network Analyst**. Select **Flow/Display Arrows**



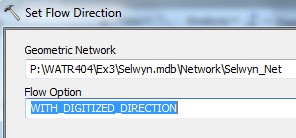
and you’ll see lots of blobs appear. This means that the network edges don’t know in what direction water flows on them.



Use the Search Function to select the Set Flow Direction function



Select the **Selwyn\_Net** for the Flow Direction and **WITH\_DIGITIZED\_DIRECTION** for the flow direction. This means that the flow will go in the direction of the nodes that make up the river lines and these are in a sequence that points downstream.

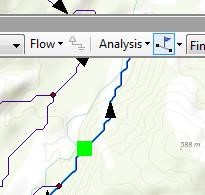
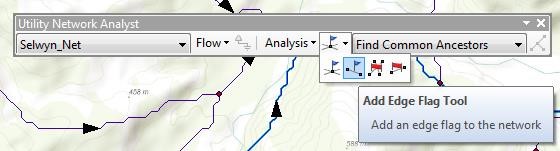
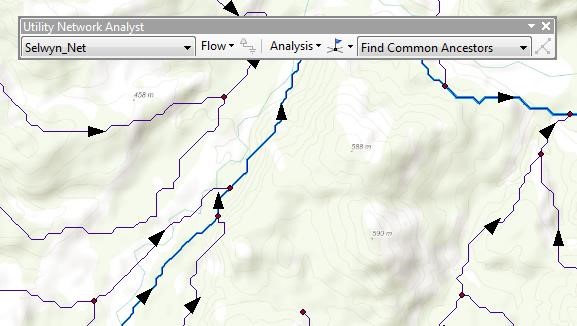


Now you can see that the network knows in what direction that the water flows. Pretty cool!!

Now let’s set an

**Edge Flag**

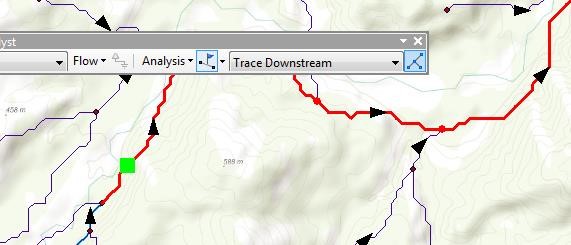
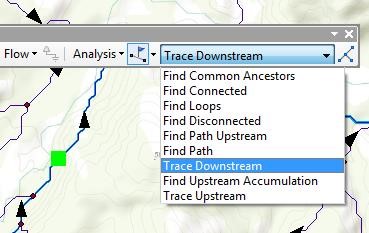
on a network edge.



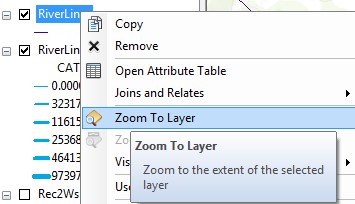
Now, let’s do a Trace Downstream from this location

Click on the

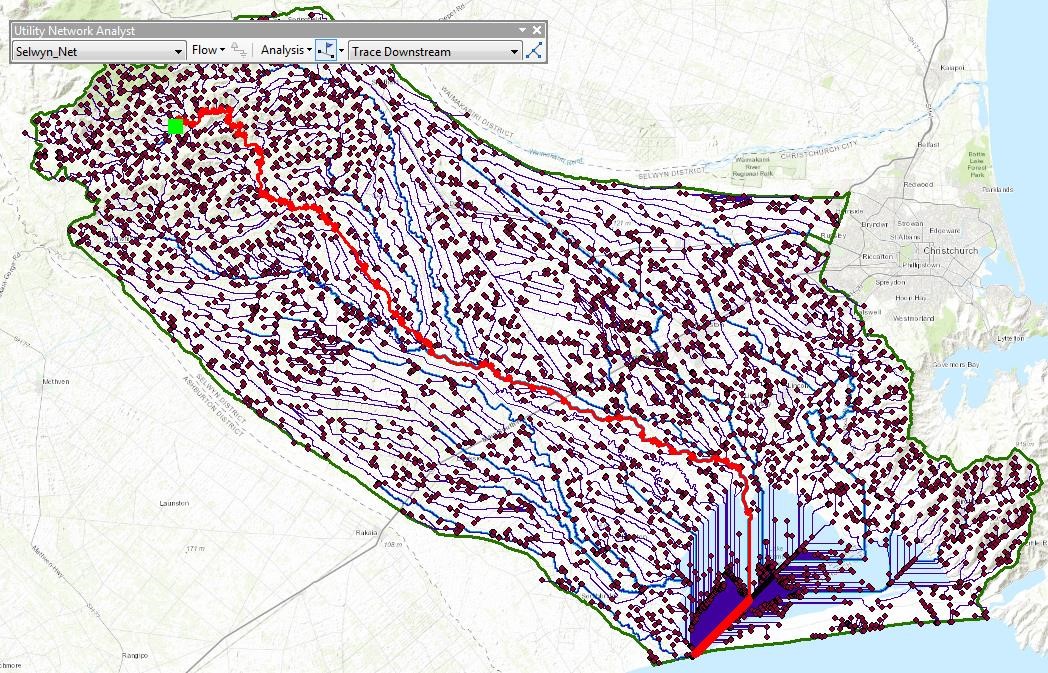
and now you’ll see a path showing how the flow goes downstream.



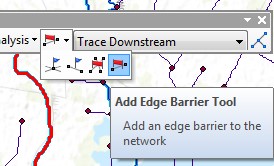
If you **Zoom to Layer** extent and turn off the Show Arrows option,



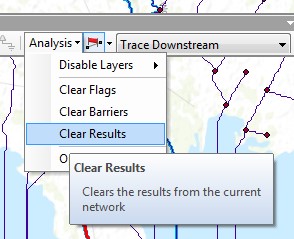
And you’ll see a flow path all the way down to the sea.



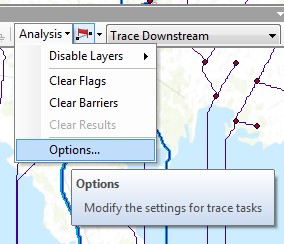
For assessment of the impact on Lake Ellesmere/Te Waihora, we don’t want the distance all the way to the sea but rather to the shoreline of the lake, so let’s zoom in there, and set a Barrier Barrier Tool, which is a red X that is a bit hard to see in this display.



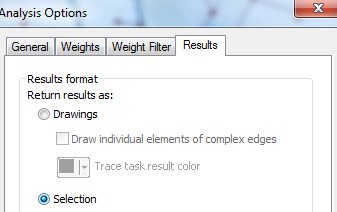
We select **Analysis/Clear Results** and it’s a bit easier to see



Now, let’s select **Analysis/Options** and

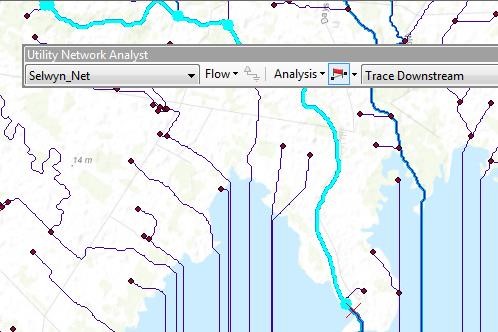


And return **Results** as **Selection** rather than Drawings

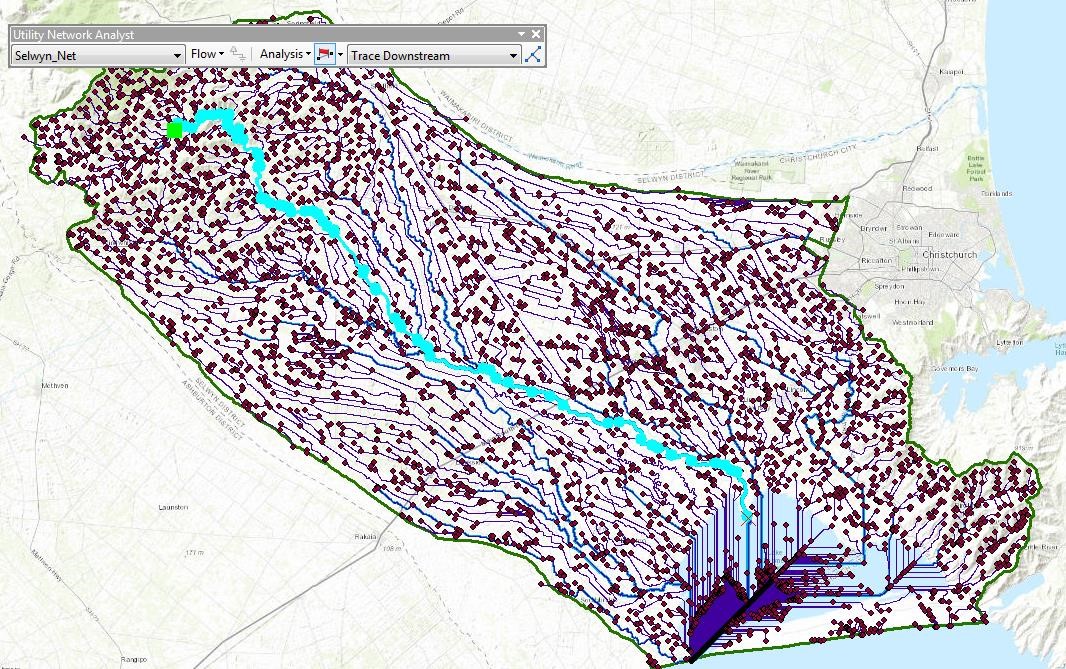


And now let’s do our Trace Downstream again

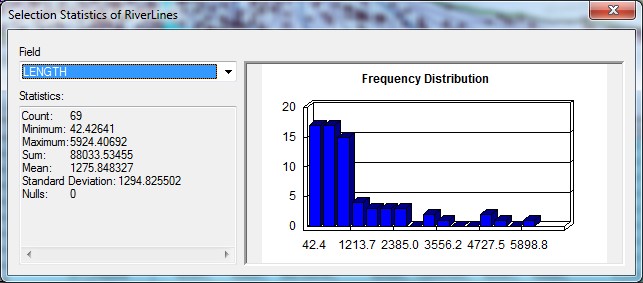
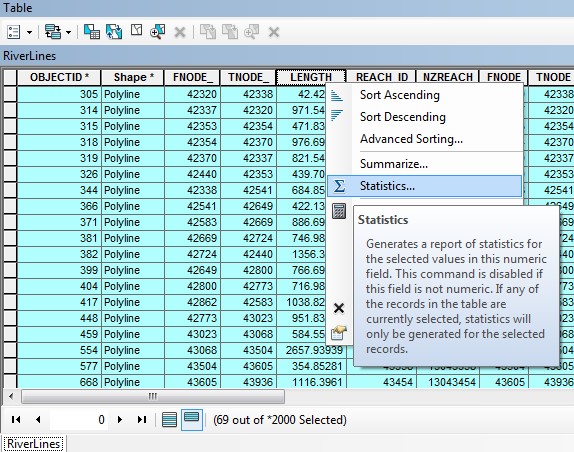
And you’ll see that the riverlines are now selected in the display



And if you Zoom to Layer again, you can see a very nice tracing of the flow path from our selected reach to Te Waihora.



If you open the Attribute Table, hit the Selected Record button at the bottom of the screen, and then calculate the Statistics on the Length attribute



In this the Sum value is 88033 meters or 88.033 Km. This means that water and contaminants flowing along this reach will have to travel about 88 Km to get to Te Waihora. As contaminants travel in water, they are transformed – the concentration of bacteria, for example, drops significantly with time (say 80% loss in 1 days travel time). So travel distance and time are important indicators of the likelihood of deleterious downstream impacts from upstream contamination.

*To be turned in: Make a nice map of the Selwyn catchment with a pathway from one of the upstream reaches to Lake Ellesmere/Te Waihora shown as selected. What is the length of this flow path in Km?*

**Summary of Items to be Turned in:**

1. *A screen capture of the header of your WaterML response for flow along with the first couple of data values, as shown above. What is the time period of your data request (from date, to date)? What are the units of the flow data? What was the mean daily flow of the Mangatainoka River at Pahiatua Town Bridge on 5 March 2018?*
2. *A screen capture of the header of your WaterML response for E. Coli along with the first couple of data values, as shown above. What is the Feature of Interest, Observed Property, Procedure of Measurement and Temporal Domain of the Result? What is the time period of your data request (from date, to date)? What are the units of the E. Coli data?*
3. *Plot a nice time series chart of these data and find the critical percentile values of these data that correspond to the New Zealand coliform data standards. What quality level do these data represent?*
4. *Make a nice map of the Selwyn catchment with a pathway from one of the upstream reaches to Lake Ellesmere/Te Waihora shown as selected. What is the length of this flow path in Km?*