Full Protocol for GeoChemNet Usage

1. Make the network look right.

* Element choices. If the network has too many element nodes right now making it hard to interpret, consider using domain knowledge to start with a smaller subset. As you understand your geochemical data, you can add more elements to the network in later stages of your exploration.
* Threshold choices. If the network has too many soil samples, and is taking more than a minute to run, consider lowering the percentile threshold. This means you will understand behaviors in the most anomalous first, and can always include more samples later by increasing the threshold.
* Threshold method. The default “scaling” based thresholding respects original elemental distributions most. If you find nodes that have abnormally many samples around them, that element probably has a very skewed distribution. Consider removing it, or switching to “percentile” based thresholding.
* Omission of univariates.
* Hyperparameter choices. “k multiplier” affects the ideal distance between nodes. “Weight exponent” affects the relative strength of springs between low and high concentrations. If low anomalous nodes seem to be very far from the main part of the network, consider lowering the “k multiplier” and increasing “weight exponent”. Might need a bit of tinkering around.
* Adjusting x and y axes limits. The network plots in a network space where the x-axis and y-axis have little numerical meaning. Change the default x and y axes limits to comfortably fit the size of the network.

2. Look for groups in data.

* Find large groups of samples that lie together in the network space. These are points that behave geochemically similarly.
* Utilize the selection widget to outline points of interest. Try to find what anomalies led to the creation of these groups. For example, if a group of points is all connected to Cu, use the textbox to identify if it is a “high Cu” or a “low Cu” anomaly. Do this to all connected elements to find multidimensional anomalies, eg, determining the group is “high Cu, low Ti, low Al”

3. Color groups.

* Utilize the dropdown menu widget to change the color of the samples. Here you can find other elemental relations to these anomalies.
* If available, color by other attributes of importance that exist in your meta data. If you have an age or geologic unit, you can observe where these properties lie on the geochemical network space.

4. Consider geospatial trends.

* If available, use the coloring and selection tool to observe anomalies in the geochemical and geospatial realms. Look for geospatial groupings of points that color and select similarly, given that these nonrandom clusters probably reflect locations of interest.

5. Rinse and repeat.

* Repeat steps 2-4 until all major groups in the network have been explored.
* Consider repeating steps 1-4 using a network that includes more points. Consider increasing your percentile threshold, hence defining more things as anomalous, to understand a bigger proportion of your data. If you have removed elements, consider adding more elements. Note, all anomalous groups found in this smaller starting point will also appear in more expanded network explorations. Hence, further explorations will add to known geochemical behaviors, or uncover new ones.
* Consider repeating steps 1-4 with a different seed. When creating the networks and reaching equilibrium of pushes and pulls of nodes and springs, randomness is involved. By using a new seed, you will explore a new equilibrium position of nodes. Groups of points and their qualities will stay consistent, but perhaps certain equilibrium positions may highlight certain groupings better.