First of all, we joined data for all days for all hours. From that we are extracting average delay and average connectedness(with respect to maximum number of connections) for pairs of staions. Then, we took average delay for each station, average connectedness and average distance between stations. With all that information, we can do analysis.

For stations:

As we can expect, statiton with smaller distances are more likely to be connected(between themeselves). From the plots delay vs distance vs connectedness we can see that time delay doesn't affect connectedness and distances(oscillation on 2D scatter plots are small) and connectedness has the biggest role in figures.

For pairs of stations:

It's interesting that delay vs distance is pretty random plot(?).

After cross-correlation of time delays between pairs of stations, we take argmax and se how much we need to shift signals. Smallest shift between all pairs of stations is 593(hours).