Source properties of induced seismicity at Rot/Nga

Chet Hopp[1]

[1]Victoria University of Wellington

November 11, 2018

Multiplet cluster analysis

Many of the 1171 template events have highly similar waveforms, implying similar locations and source mechanisms as well as similar detections. As treating these near-repeating events separately added little to our analysis, we employed a heirarchical clustering technique to the group template events based on waveform similarity. We will refer to these groups as multiplets.

To do this, we created a matrix of pair-wise, normalized cross-correlation coefficients for the northern and southern clusters separately. The corresponding distance matrix is the complement of the correlation matrix (i.e. 1 - cross-correlation coefficient) (see Figures **??**, **??**) and is symmetric about the diagonal. Each cell represents the pair-wise distance between two events. As the matrices are symmetric, row n and column n are identical and the cell where they meet, on the diagonal, represents a perfect correlation of 1.0 as well as a distance of 0.0. To cluster the distance matrix into multiplets we used a single-link hierarchical clustering algorithm, which can be visualized using a dendrogram. The rows and columns of the matrices shown in Figures **??** and **??** are organized via the dendrogram shown above and to the left of the matrix.

We defined a multiplet as any group of links which fell below a distance of 0.35 (or above a correlation coefficient of 0.65) (Figures **??**, **??**). We chose 0.35 as a distance cutoff based on visual inspection of the dendrograms. We found this was a more reliable method of distinguishing between multiplets than an automated method based on inconsistency. As seen in Figures **??** and **??**, many events were not grouped into multiplets because they were not linked with any other events at a value below the distance cutoff. These events are referred to as singletons. It is important to note that, while they do not cluster with other events, some singletons do detect significant seismicity.

Cluster Analysis

At northern Ngatamariki, the cluster analysis described above yielded 19 multiplets with sizes of between 2 and 54 events. 151 events were included in clusters and there were 281 singleton events from the 432 template events in the northern cluster. Figure **??** shows the full, clustered distance matrix for northern Ngatamariki templates. Numbers are assigned to singletons and multiplets as a way of distinguishing them throughout this text. The first singleton on the left of the dendrogram is assigned number 1, then number 2, and so on from left to right, with multiplets being assigned only one integer for all the events within. Multiplet 95 is the largest in northern Ngatamariki and is a focus of much of the following discussion. The waveforms for this particular cluster are shown in Figure **??** on the vertical channel at station WPRZ, located north of the northern injection wells.

For southern Ngatamariki, the same analysis yielded 42 multiplets with sizes between 2 and 166 events. Of the 739 template events in southern Ngatamariki, shown in Figure **??**, 315 were clustered and 424 were singletons. Multiplet 150 is the largest multiplet in southern Ngatamariki and is labeled in Figure **??**. Figure

?? shows the waveforms for multiplet 150 on the vertical channel of station NS14, located near the well pad for NM06 and NM10.

Multiplets are diverse not only in their average locations and source parameters, but also in their internal structure. For example, in northern Ngatamariki, the waveforms for events within multiplet 95 (Figure **??**) could feasibly be clustered further into smaller, more coherent multiplets. Partially, this is an artifact of the threshold we've chosen for distinguishing between events falling into different multiplets. If we look again at multiplet 95, we see that there are at least 3 sub-multiplets. The same could be said of multiplet 96 albeit at a higher distance threshold. This is easily seen in the dendrogram at the top of Figure **??**. Multiplet 95 contains 54 families, corresponding to 54 template events, but within those columns of the dendrogram there is yet more structure which represents further divisions in the data which were not picked up in the initial clustering.

The single-link dendrograms are constructed by first finding the two best correlated events. We'll call these events 1 and 2 and they have the lowest distance between them among all possible events. When visualizing the dendrogram, this distance corresponds to the distance between the bottom of the dendrogram and the line connecting the two events, which is called a 'link'. Next, all remaining pairs are searched for the pair with the lowest distance. This third event, event 3, is then linked to the original events in the same way as event 1 and event 2. Event 4 is linked to events 1, 2 and 3 with a link equal to the lowest distance between it and any of the original three events. This process continues iteratively until all events in the matrix have been linked. We chose this 'single-link' algorithm because it has the effect of creating large clusters of events. For example, an event will be linked to a preexisting cluster without having to correlate highly with each event within this cluster, instead having to correlate highly with only one. This is desirable because we are interested in looking at how events with similar source mechanisms and location vary in time within an actively exploited reservoir.

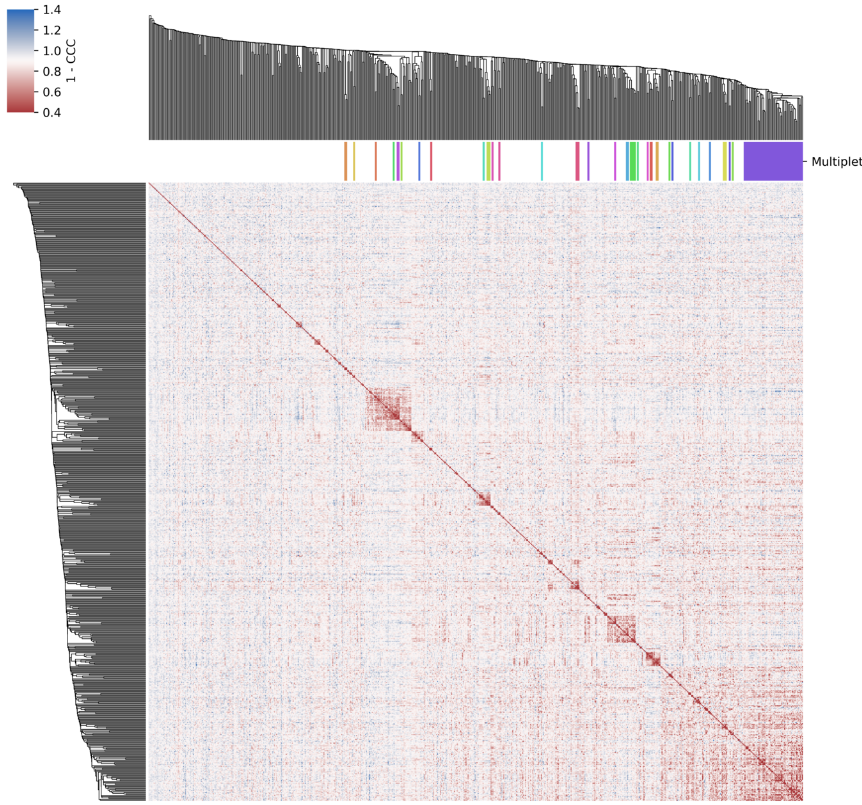

Figure 1: Clustered distance matrix for sorthern Ngatamariki, which is symmetrical about the upper left-to-lower right diagonal. Cells are colored by distance. Red signifies a high correlation coefficient, blue the opposite. Each row or column represents a single seismic event (e.g. the first row and first column are identical, as are the second row and second column, etc...). Above the matrix, multiplets are shown as colored tabs above the columns which pertain to each. Multiplets which are referenced in this text are labeled with their corresponding multiplet number, which is essentially arbitrary and used only for distinguishing between multiplets. The single-link dendrogram is plotted at the top and left which shows the structure of the distance matrix which was used in the clustering analysis.

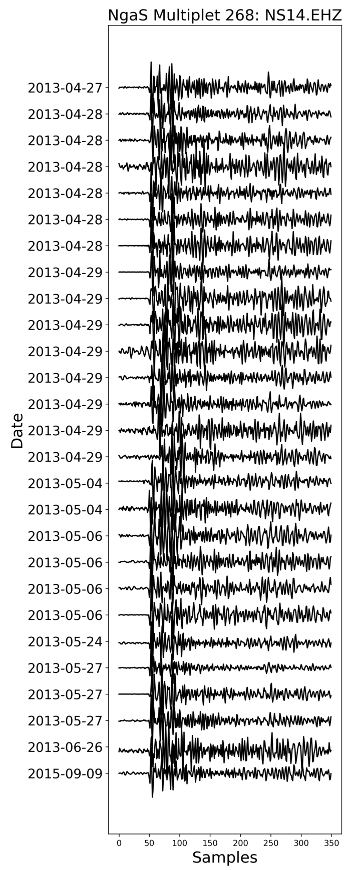

Figure 2: Waveforms for southern multiplet 268 on the vertical channel of borehole station NS14 (depth 202 m). Data are sampled at 100 Hz (i.e. 3.5 seconds) and aligned on the automatic P-pick from the GNS catalog. Waveforms proceed chronologically from top to bottom.

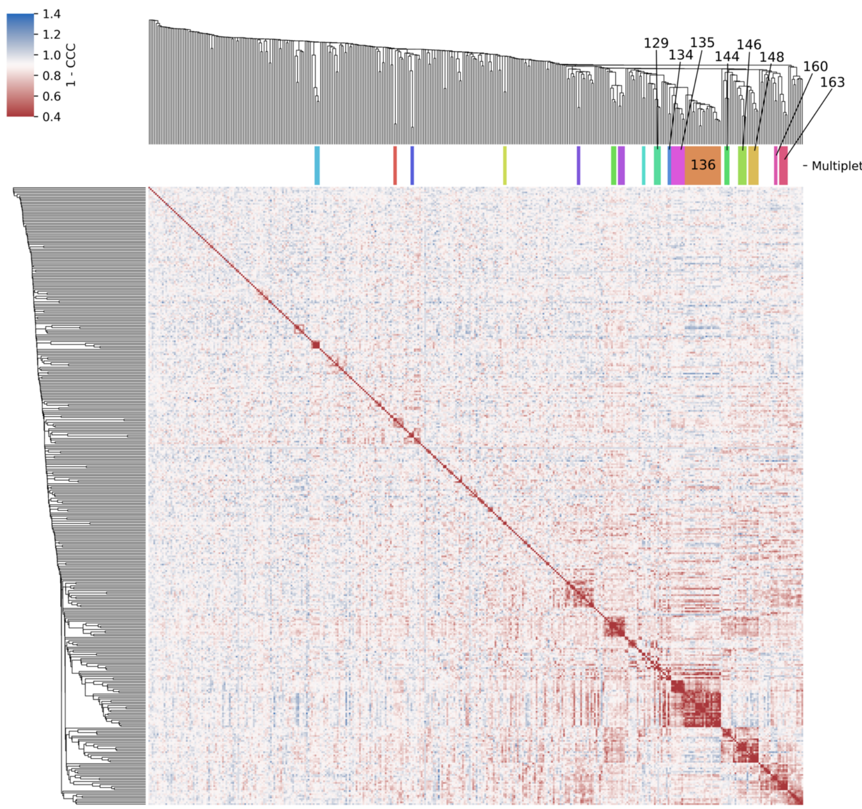

Figure 3: Clustered distance matrix for northern Ngatamariki, which is symmetrical about the upper left-to-lower right diagonal. Cells are colored by distance. Red signifies a high correlation coefficient, blue the opposite. Each row or column represents a single seismic event (e.g. the first row and first column are identical, as are the second row and second column, etc. . .). Above the matrix, multiplets are shown as colored tabs above the columns which pertain to each. Multiplets which are referenced in this text are labeled with their corresponding multiplet number, which is essentially arbitrary and used only for distinguishing between multiplets. The single-link dendrogram is plotted at the top and left which shows the structure of the distance matrix which was used in the clustering analysis.

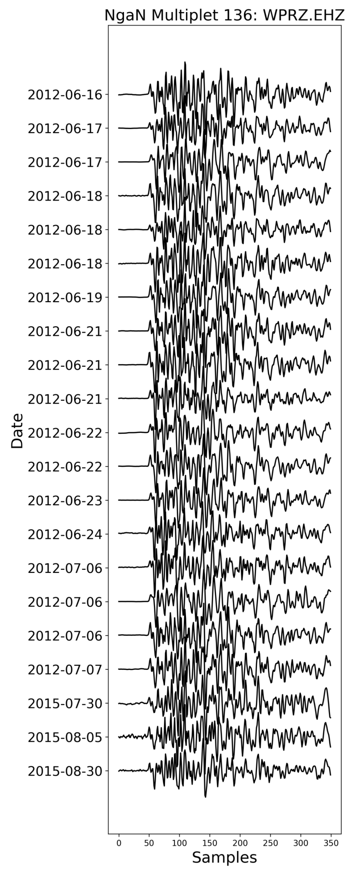

Figure 4: Waveforms for northern multiplet 136 on the vertical channel of station WPRZ. Data are sampled at 100 Hz (i.e. 3.5 seconds) and aligned on the automatic P-pick from the GNS catalog. Waveforms proceed chronologically from top to bottom.

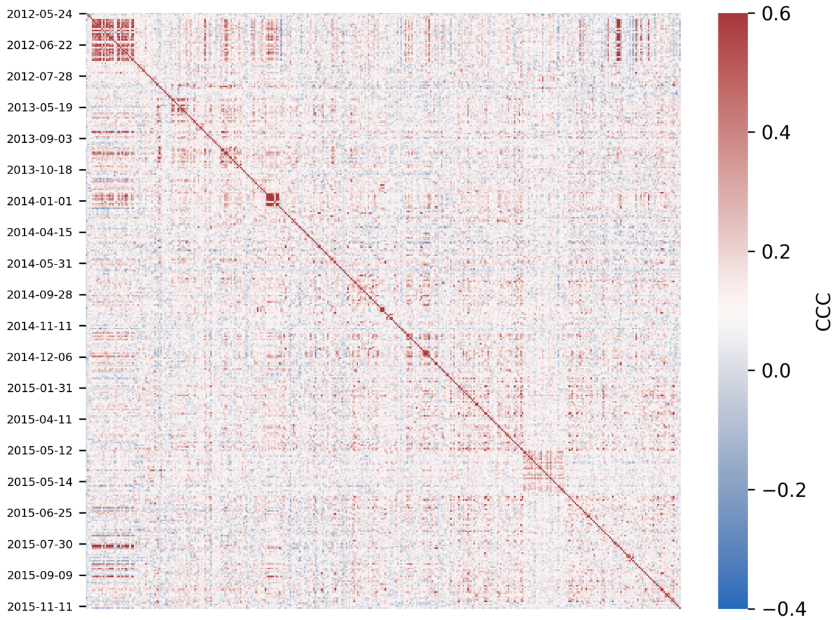

Figure 5: This is a caption

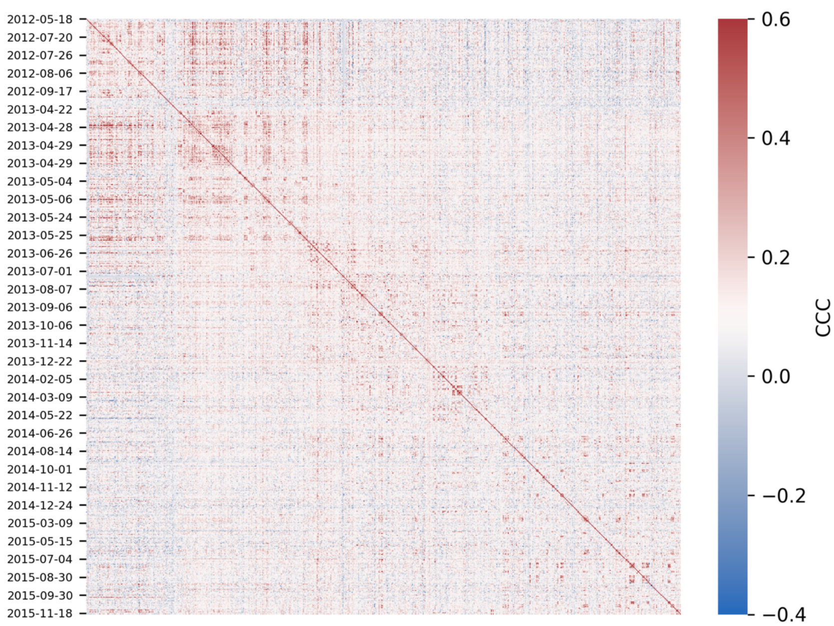

Figure 6: This is a caption