PROCEEDINGS, 51st Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 9-11, 2026
SGP-TR-230

Spatial Clustering and Analysis of Microearthquakes in Tiwi Geothermal Field through OPTICS Algorithm

Al Christian R. Gobres[1] and Hans Jake F. Felosopo[1]

[1]Philippine Geothermal Production Company, Inc., 14th Floor, 6750 Building Ayala Avenue, Makati, Manila, Philippines

AGobres@pgpc.com.ph

Keywords: Microearthquakes, OPTICS, clustering, machine learning, Tiwi Geothermal Field

ABSTRACT

Microearthquakes (MEQs), also known as microseismic events, are weak seismicity with magnitudes below 3.0 that do not present any risk to nearby communities or infrastructures. Their occurrence is induced and thus inherent to geothermal field operations, playing an important role in monitoring changes in reservoir stress fields. Events typically occur when conditions in pre-existing faults or fracture networks change due to fluid extraction and brine injection, causing stress buildup that eventually reaches a threshold causing rock failure or slippage, releasing stored strain energy as MEQs. Mapping and analyzing the spatial distribution of these events can contribute to the development of injection management strategies especially when integrated with injection tracer test data. The same is also used to confirm the viability of an area as a drilling target. Spatial clustering of MEQs is employed to uncover patterns and trends, ultimately deriving valuable insights related to reservoir structures and dynamics. Currently, conventional methods like grid-based clustering calculate event density within fixed grids and has been proven to be effective in identifying anomalies or dense clusters. However, this approach struggles to distinguish nested clusters or account for the depth component of the events during clustering. To address these limitations, unsupervised machine learning algorithms have gained prominence in identifying significant patterns of arbitrary shapes within large datasets. In this study, the OPTICS (Ordering Points To Identify the Clustering Structure) algorithm, a density-based clustering method, is applied to a 3D spatial dataset of MEQ events from the Tiwi Geothermal Field between 2012 and 2017. This research aims to examine the impact of input parameters on clustering solutions and identify regions of interest for detailed analysis.

1. INTRODUCTION

1.1 Background

1.1.1 Microearthquakes in Geothermal Field

Microearthquakes (MEQs), also referred to as microseisms or microseismic events, are typically very low-intensity earthquakes occurring within the geothermal reservoir rocks of active geothermal fields. These are generally induced seismicity having magnitudes below 3.0, commonly associated with steam production and reinjection of separated geothermal fluids back into the reservoir.

Majority of MEQs are induced by cumulative changes in the stress field, which result from variations in temperature and pressure due to field injection and production activities. As a result, MEQs are commonly inherent in operating geothermal fields. The most frequent triggering mechanisms is cooling effect from fluid injection or marginal recharge, wherein high temperature reservoir rocks are stimulated through interconnected deep fractures, generating temperature contrasts. A thermoelastic (Mossop and Segall, 1997) triggering mechanism, such as thermal contraction of reservoir rock in response to injection of cooler fluids, causes fracture apertures to expand, allowing seismic slippage along the previously locked fracture surfaces (Figure 1).

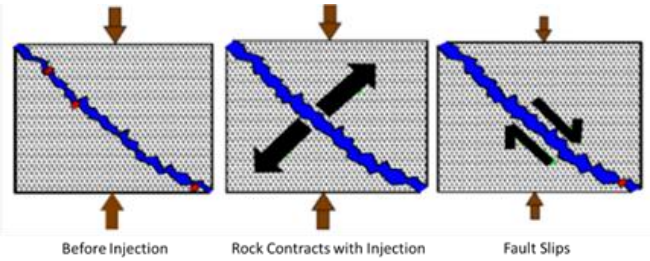

Figure 1: Schematic illustrating MEQs induced from thermal contraction mechanism (Marcuap and Villamer, 2019)

High-pressure fluid injection is another potential trigger. Additionally, mass extraction and thermal contraction may result in reservoir compaction, which can cause ground subsidence and corresponding MEQs. The reduction in reservoir pore pressure caused by geothermal fluid extraction generates pressure differentials that lead to seismic activity.

Since MEQs in geothermal fields are induced, activities such as fluid injection can map seismic patterns that can be used to infer the movement of injected fluids, valuable for optimizing injection strategies. MEQ data can also support operational decisions, such as

whether to shut in an injection well or redirect injection to another location, particularly when supported by geochemical and well surveillance indicating a negative impact on production zones. Additionally, these seismic patterns can be correlated with reservoir structures—such as conduit faults, fault boundaries, and permeability zones associated with upflow—to enhance understanding of the geothermal resource. During the exploitation phase, the distribution of MEQs may also help delineate the extent of the production area and the base of the reservoir.

Regular monitoring of seismicity levels in operating geothermal fields is submitted to the Department of Environment and Natural Resources (DENR) as required, with the objective of monitoring the impact of seismic activity on surface facilities and local communities. The effect of Make-Up Well (MUW) drilling on local seismicity is also closely monitored and in PGPC's experience, these have been shown to be minimal, posing no threat to surface facilities or nearby communities. Future drilling operations may be impacted in areas where stress fields have been disrupted, or where reservoir compaction and subsidence have occurred, increasing the risk of formation collapse.

1.1.2 Seismic Clustering

Clustering is a technique used in data analysis and machine learning to group a set of objects in such a way that objects in the same group (called a cluster) are more similar to each other than to those in other groups. This approach is particularly useful when working with large spatio-temporal datasets. In geothermal studies, microearthquakes are characterized by their hypocenters—data points defined by their X and Y coordinates, depth (Z) component, and time of occurrence, serving as key indicators of subsurface seismic activity.

The hypocenters of microearthquakes represent the precise locations within geothermal reservoir rocks where seismic energy is released due to rupture. The primary goal of clustering in this context is to identify and map regions of dense seismicity—sometimes referred to as MEQ anomalies—within the geothermal field typically observed over extended time periods. Given the time-dependent nature of this approach, MEQs can be correlated with surface facility operations, allowing researchers to identify how certain surface manipulations or interventions may trigger localized seismic activity.

By mapping the shape and extent of the dense seismic clusters, insights regarding subsurface reservoir structures and associated permeability pathways can be inferred. Both fluid withdrawal and injection can alter the subsurface stress field. Injected fluids, in particular, act as a medium that travels through fractures and faults, often triggering seismic events along their path. Ground subsidence and reservoir compaction from fluid withdrawal occurs where the production zones are. As a result, this method can aid in delineating the extent and base of the permeable reservoir, identifying fault conduits, and mapping bounding geological structures that influence fluid flow and seismic activity. Furthermore, seismic clustering analysis can help distinguish between different reservoir and surface processes that induce microearthquakes. For example, deeper clusters may be linked to injection activities, while shallower ones may correlate with production operations, offering valuable information about the dynamics of geothermal systems.

Several types of clustering techniques include hierarchical clustering methods, partitional clustering, grid-based clustering, and density-based clustering, varying based on the similarity measures each algorithm considers in forming clusters. The first two techniques mentioned, however, pose a disadvantage in spatial clustering as these methods are sensitive to noise and center point, and can only detect spherical clusters (V, 2013).

1.1.3 Grid-based Clustering

Grid-based clustering is a widely recognized method among conventional clustering techniques. It offers a practical and efficient approach for analyzing spatial data, particularly in geothermal studies where microearthquake distributions are of interest. One of its key advantages lies in its scalability and computational efficiency, as it partitions the data space into a finite number of cells (or grids) with fixed sizes, and evaluates density based on the number of points within each cell. This method simplifies clustering by avoiding complex distance calculations and is especially well-suited for large datasets. Additionally, it aligns naturally with geospatial data, making it ideal for identifying zones of concentrated seismic activity. Effective noise handling, low-density or sparse cells (outlier points) can be easily filtered out, helping to reduce noise and highlight meaningful patterns. Grid-based clustering does not require the user to predefine the number of clusters, making it more fl*exi*ble in exploratory analysis. It has also deterministic results, unlike some clustering algorithms that rely on random initialization (e.g., k-means), grid-based clustering produces consistent results given the same parameters.

However, grid-based clustering also comes with limitations. The choice of grid resolution is critical—too coarse a grid may overlook smaller clusters, while too fine a grid can fragment meaningful groupings and increase computational overhead. Moreover, the rigid structure of grid cells may not accurately capture the irregular shapes of natural phenomena like fault zones or fluid pathways. The method also relies heavily on predefined parameters such as grid size and density thresholds, which may not generalize well across different datasets. Lastly, while effective in low-dimensional spaces, grid-based clustering can struggle with higher-dimensional data, where its interpretability and performance may decline. A key limitation of grid-based clustering when applied to 3D spatial data—especially in geothermal studies involving the Z or depth component—is its infle*xi*bility in capturing vertical variations in seismic activity. Grid-based clustering typically uses uniform cell sizes across all dimensions. In geothermal systems, seismic activity can vary significantly with depth, and a fixed grid size may fail to capture subtle but important vertical patterns. It also prone to loss of vertical detail, such that if the grid is too coarse in the Z direction, microearthquake clusters that are vertically stratified (e.g., along fault planes or fluid pathways) may be merged or missed entirely. Conversely, a very fine grid in depth may lead to fragmentation of clusters, making interpretation difficult. Geological structures are often irregular, natural subsurface features like faults, fractures, and reservoirs are not aligned with regular grid boundaries. This misalignment can cause grid-based clustering to misrepresent the true shape and extent of seismic zones.

Grid-based clustering is simple and powerful for spatial data in 2D or 3D. But as more dimensions are added, it becomes harder to manage, interpret, and extract useful insights. For high-dimensional data, other clustering methods like DBSCAN, hierarchical clustering, or model-based approaches may be more suitable.

1.1.4 Density-based Clustering

Density-based clustering techniques group objects into clusters by checking that each object in a 'cluster' contains at least a minimum number of objects, *min_pts*, within the neighborhood of a given radius, ε. *Min_pts* and ε are the two required input parameters in a density-based clustering algorithm. A cluster contains a 'core object' and 'border objects'. An object p is classified as a core object if the number of objects within its ε neighborhood is greater than or equal to the minimum points requirement. The point is considered as a border object, on the other hand, if and only if the number of points in its ε -neighborhood is less than *min_pts* and the considered point is contained in the ε-neighborhood of a core point. Points that are neither core nor border points are considered as 'noise points' (Ankerst, 1999).

DBSCAN, or Density Based Spatial Clustering of Applications with Noise, was introduced by Ester et al. (1996) which uses the above technique in clustering objects. The algorithm works by checking the ε-neighborhood of each point in the data set. If the point p classifies as a core object based on its definition, a new cluster C is created containing the objects in its neighborhood. The ε-neighborhood of other points q contained in C are checked one-by-one. If the ε-neighborhood points of q contain more than *min_pts*, the neighbors of q not yet contained in C are added to the cluster. These steps are repeated for their respective ε-neighborhood points until no new points can be added to the current cluster C (Ankerst, 1999). This algorithm is useful in cluster analysis of 3-D spatial data as the cluster geometry is not predefined and density-based clusters of any shape can be identified (Piegari, 2022).

One of the major drawbacks of DBSCAN is not being able to distinguish regions with different densities (Piegari, Herrmann, & Marzocchi, 2022). Ankerst (1999) proposed OPTICS (Ordering Points To Identify Clustering Structure), an extended-DBSCAN algorithm. OPTICS works by clustering the data points at an infinite number of distance parameters simultaneously. Objects which are density-reachable with respect to the lowest ε are processed first, ensuring clusters with higher density (same *min_pts* at lower ε) are prioritized. The order in which the data points are processed and information needed for the extended DBSCAN are stored to provide consistent result. These additional information are called core distance, d_C , and reachability distance, d_R . Core distance of a point p is defined as the minimum neighborhood distance such that p is a core point i.e. number of points within its ε-neighborhood is at least *min_pts*. Reachability distance between two points p and q will only be defined if p is a core point, and d_R is the greater value between its core distance and distance between p and q. Compared to DBSCAN, OPTICS does not require ε as an input parameter as it spans its search radius from zero to infinity, although it can be set to a very large number to save computation time (Piegari, 2022).

OPTICS works by iterating through each unprocessed point in the data set, identifying the density clustering structure given by reachability distances. Points q within its neighborhood are identified. If q is a core point, the reachability distance of its unprocessed neighbors are updated. These points are added to the priority queue based on their respective reachability distances, where the points closer in density, corresponding to lower reachability distances, are processed first. Reachability distances are stored with a corresponding order, resulting to a reachability plot.

OPTICS does not directly output the clustering solution, but instead provides an ordering of points represented as a reachability plot. Clusters correspond to 'valleys' in the plot. The smaller d_R , the denser the cluster. The clusters are separated by 'peaks' in the plot: the higher the peak, the more separated the clusters are. Smaller valleys can e*xi*st in larger valleys, which represent nested clusters. It is important to take note that the cluster order is not fully deterministic because points with the same distance may be processed in any order. Thus, different runs may yield different results that nevertheless correspond to a highly similar cluster structure (Schubert & Gertz, 2018).

Clusters from the reachability plot are extracted using several methods: by visual inspection, using significant ma*xi*ma, inflex*i*on points and the ξ-steep method. These methods are all based on how the 'valleys' are defined, and are attempts to automate the valley-detection process. The ξ-steep method was described by Ankerst (1999) for identifying potential start-of-cluster and end-of-cluster regions first, then combine matching regions into clusters. OPTICS define the beginning of a cluster by downward steep points i.e. points are succeeded by a point 1- ξ lower reachability, and the end of a cluster are signaled by upward steep points i.e. points are succeeded by a point 1- ξ higher. The number of points within the cluster area should be greater than or equal to *min_pts*.

The number of clusters generated by unsupervised clustering algorithms such as OPTICS is determined based on cluster stability. A clustering solution is deemed stable if similar results are consistently obtained across a range of input parameter settings.

The problem is depending on the spatial distribution, even small changes in the input parameters can lead to very different cluster solutions. (Piegari et al., 2022) explored the challenges in the calibration of the procedure to obtain stable cluster solution. This was also considered in this study.

1.2 Objectives

This study was aimed to identify MEQ clusters from the 2012–2017 MEQ dataset of the Tiwi geothermal field using the density-based clustering algorithm, OPTICS. The sensitivity of the clustering solutions was also explored by varying the input parameters. The results of the OPTICS-based clustering were also compared with those obtained through grid-based clustering on the same dataset. Preliminary insights were then derived from the solutions and correlated with subsurface fluid injection and production activities based on the identified seismic patterns.

1.3 Scope and Limitations

This study focused on the identification and analysis of MEQ clusters or zones of dense seismicity within the Tiwi geothermal field using microseismic data collected from 2012 to 2017. The study focuses on the application of OPTICS algorithm to detect patterns of seismicity and delineate zones of concentrated activity. The clustering results of OPTICS algorithm was compared only with the clustering results obtained from grid-based technique. Furthermore, the analysis is confined to geothermal context, aiming to understand subsurface seismic patterns and derive initial insights related to subsurface fluid injection and production processes by interpreting the spatial distribution of generated MEQ clusters.

The MEQ dataset and analysis is limited to a specific location and time frame. This paper discusses the MEQs in the Tiwi geothermal field, within the 2012–2017 monitoring period, more than three decades after production commenced in 1979. This may restrict the generalizability of findings to other geothermal fields or time periods. Grid-based clustering may oversimplify spatial patterns, particularly in the depth (Z) dimension, due to its reliance on fixed cell sizes and rigid boundaries. The accuracy of clustering results is influenced by the quality and resolution of the seismic data, including uncertainties in hypocenter depth estimation. OPTICS and grid-based clustering methods may yield different interpretations depending on parameter settings, such as density thresholds and grid resolution, which require careful tuning. Moreover, data gathering limitations of the MEQ seismic network deployed in the field, such as array aperture or sensor coverage, data scatter or resolution, variability in manual picking accuracy (RMS), depth of investigation or accuracy of hypocenter estimations particularly in depth, may affect the precision of clustering results. The study provides only preliminary insights into subsurface processes; a more comprehensive geomechanical or hydrogeological analysis would be needed to fully understand the implications of MEQ clustering. Additionally, the study does not incorporate other geological or geomechanical data that could provide more profound insights into subsurface processes.

2 METHODOLOGY

2.1 Data Gathering and Processing

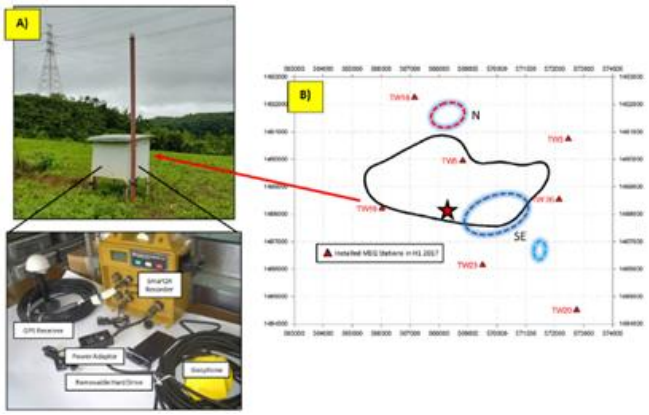

Figure 2: A) MEQ monitoring station setup in the field, housed within an enclosure that contains the instrument and its accessories. B) Tiwi MEQ monitoring station map in H1 2017. The black boundary outlines the production area, while the dashed lines in indicate the field's injection zones. The red star marks the approximate location of the southern upflow.

The current MEQ monitoring network in Tiwi was installed in mid-2012 primarily to monitor the response of the reservoir to the brine diversion program (Calibugan et al, 2015) from the hot brine injection system in the north (red dashed lines) towards the southeast of the field (blue dashed lines). Five years of continuous monitoring showed that the highest rate of microseismic events is in the southern part of the field where an upflow (red star) is located (west of the southeast hot brine injection system). These microseisms are seen as reactions to pressure and temperature changes in the reservoir mainly resulting from the regular injection of fluids that are cooler than the reservoir temperature.

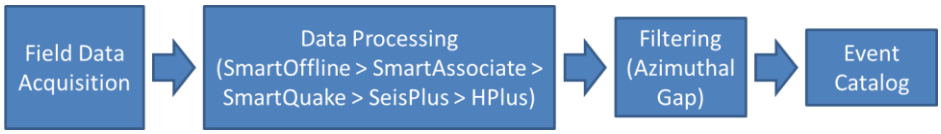

Figure 3: General workflow to derive MEQ event catalog

Field visits to each monitoring station were conducted weekly as part of the data gathering process. Data collected from the monitoring stations was extracted using the SmartOffline application and are processed to remove noise and conduct further analysis. Manual trimming of waveforms was also performed to narrow the time window, capturing the P wave, S wave, and the end of the event. After trimming, events are categorized as either local (AG ≤ 200°) or regional, followed by manual picking of the arrival of P and S waves and further processing.

The results of manual picking are reviewed for errors, and adjustments are made as needed to maintain the RMS value at ≤ 0.15. This parameter ensures that only reliable data sets proceed to the next steps. Events with RMS travel time residuals exceeding 0.15, or standard errors in latitude, longitude, and depth exceeding 1 km, are excluded. The events were prepared for filtering, with the ma*xi*mum azimuth gap covered by the monitoring stations set to ≤ 200°, ensuring that only events within the array are included in the analysis and interpretation. Any events with an angular gap between the epicenter and the monitoring station greater than 200° are considered to fall outside the array.

2.2 OPTICS Implementation

A custom program was developed in a Jupyter Notebook environment to implement the OPTICS algorithm using the scikit-learn Python library (version 0.23.1; Pedregosa et al., 2011). Aside from the two primary input parameters, *min_pts* and *xi*, the additional implementation parameters are summarized in Table 1.

Table 1: Scikit-learn OPTICS Implementation

Parameter	Definition	Value Used
max_eps	The ma*xi*mum distance between two samples for one to be considered as in the neighborhood of the other. Default value of *np.inf* will identify clusters across all scales; reducing *max_eps* will result in shorter run times.	Default (np.inf)
metric	Metric to use for distance computation	Default (Minkowski)
p	Parameter for Minkowski metric. If p = 1, equivalent to Manhattan_distance. If p = 2, equivalent to Euclidean	Default (2, Euclidean)
cluster_method	The extraction method used to extract clusters using the calculated reachability and ordering.	*xi*
predecessor_correction	Correct clusters according to the predecessors calculated by OPTICS (Schubert & Gertz, 2018).	True

Clusters were extracted using the *xi*(ξ)-method as proposed by Ankerst et al. (1999). Similar to OPTICS, the algorithm takes two main parameters: *min_pts* and *xi* (ξ) in the identification of clusters. A custom Python program was developed to generate the reachability plot and extract the clusters.

To explore the sensitivity and influence of the two parameters to the clustering solutions, the program was applied to the data set and was made to run at a range of *min_pts* (1-30, with step=1) and *xi* values (0.01-0.1, with step=0.01). The optimal range of values used for further analysis was then determined. Once the parameters were chosen, these were used to generate the reachability plot for the data set, which was then used to extract the clusters.

2.3 Grid-Based Clustering

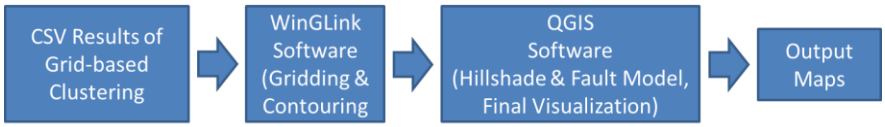

Figure 4. General workflow to derive final maps in grid-based clustering.

Gridding, contouring, and visualization procedures for the results of grid-based clustering were first performed using the WinGLink software. It provided robust tools for processing spatial data, generating contour maps, and visualizing MEQ distributions within the Tiwi geothermal field. The software's capabilities in handling geophysical datasets allowed for effective representation of subsurface patterns and clustering results, supporting the interpretation of seismic activity in relation to geothermal operations.

However, WingLink does not support the display of digital elevation models (DEMs) or topographic features directly on its maps. To overcome this limitation, the gridded outputs—such as contours and color gradient maps—were exported to QGIS for enhanced visualization. In QGIS, these results were overlaid with hillshade data derived from LiDAR and the updated fault model from Ejera, Aurelio, and Taguibao (2024), serving as the base map for better spatial correlation and geological interpretation within the Tiwi geothermal field.

Figure 5: General workflow to derive final maps in OPTICS algorithm.

2.4 Data Visualization

Clusters generated using varying parameter settings in the OPTICS algorithm were exported and transferred onto the established base map in QGIS for spatial visualization. These clusters were then compared with the results of grid-based clustering, specifically the gridded contour outputs produced in WingLink. By overlaying both sets of results in QGIS, the study was able to assess the spatial consistency and differences between the two clustering approaches, providing a clearer understanding of seismic activity patterns within the Tiwi geothermal field.

QGIS software was utilized as the primary tool for visualization and mapping throughout the study. It provided a flexible and powerful platform for displaying spatial distributions of MEQ data, generating cluster maps, and overlaying geological features relevant to the Tiwi geothermal field. The software's support for georeferenced data and customizable symbology allowed for clear representation of seismic patterns and clustering results, enhancing the interpretability of spatial relationships within the dataset. Power BI and Plotly, a Python data visualization library, were also used for data visualization to aid in the analysis of the clustering structure.

3 RESULTS AND DISCUSSION

3.1 2012-2017 Tiwi MEQ Spatial Dataset

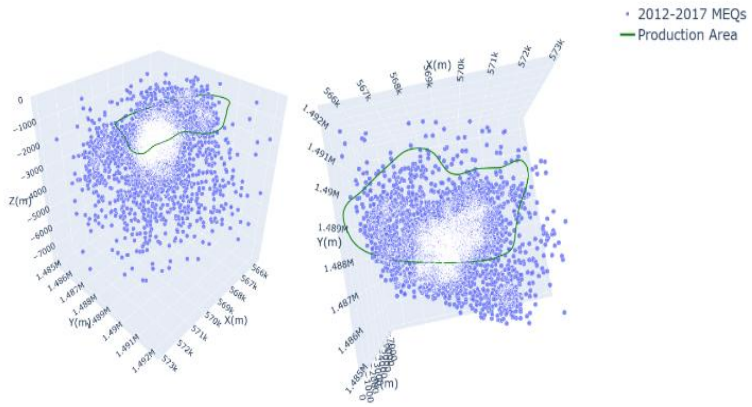

Figure 6: 2012-2017 Tiwi Microearthquake Data in 3D space

The filtered dataset is composed of 4,218 data points with corresponding X,Y and Z coordinates. A 3D visualization of the data is shown in Figure 6 with overlay of the Tiwi production boundary for reference. The X coordinates span from 559,603.4m to 579,117 m, Y coordinates from 1,478,505 m to 1,497,795 m, while the Z coordinates extends from 0 m to -7,890 m.

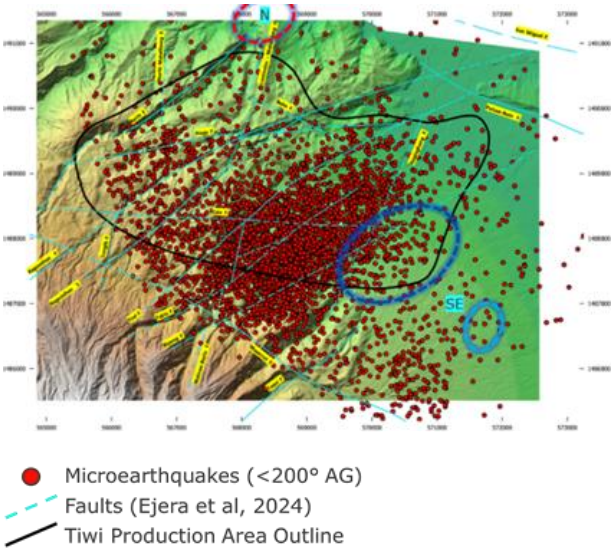

Figure 7: MEQ distribution map from May 2012 to June 2017 at the Tiwi Geothermal Field. Circles represent relative location of injection systems with respect to the production area located in the southeast (blue) and north (red).

All MEQs shown in Figure 7 are filtered events that fall within the ma*xi*mum azimuthal or angular gap of the seismic monitoring network, which is set to equal or less than 200°. This indicates that the events are occuring within the array, ensuring sufficient accuracy for reliable analysis and interpretation. Majority of MEQs in Tiwi occurred in the southern part of the field. The spatial distribution of microearthquakes in this area supports the correlation of these events with injection activities in the southeast (blue circles) where seismogenic-wells (MEQ-inducing wells) are located (Marcuap and Villamer, 2017). Epicentral distribution also provides information on where the fracture network is being stimulated. In the case of mapped structures, the intersection of Tiwi and Cale Faults suggests a fractured region where fluid flows from high pressure to low pressure (east to west) which can explain the high MEQ distribution in this area.

3.2 Grid-based Clustering

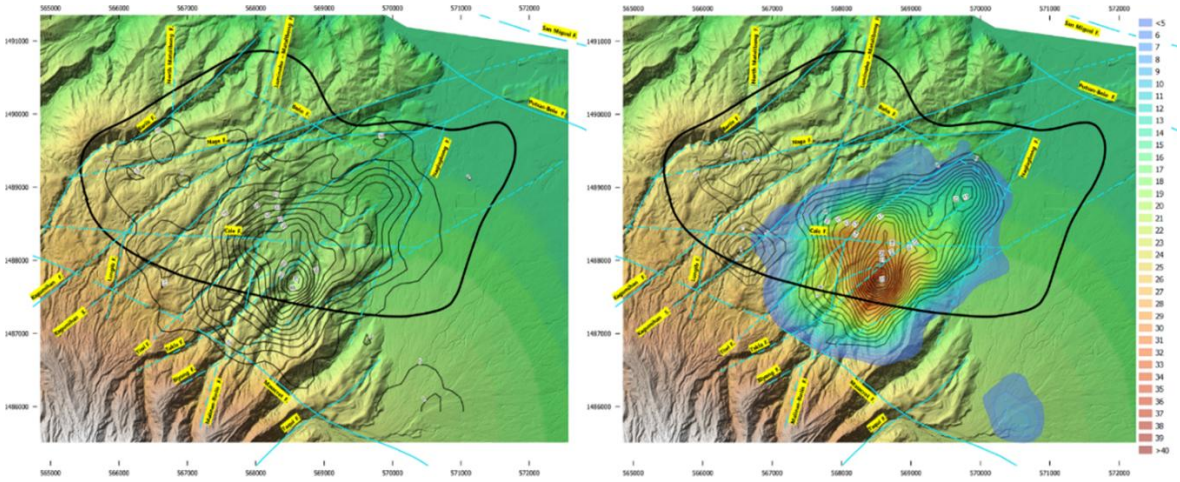

Figure 8: A) MEQ density map derived from grid-based clustering. B) MEQ density map derived from A showing shallow seismic events above 2000 mbsl (represented by contour lines) and deeper seismic events below 2000 mbsl (depicted using color gradients). Contour and color gradient values indicate number of MEQ events occurring in an area interpolated from the results of density grid analysis.

Figure 8A illustrates the MEQ density distribution across the Tiwi Geothermal Field using a grid-based clustering method. The analysis employs 250m x 250m grid cells to count MEQs per cell, generating contour maps that visualize event density. This conventional approach helps identify the spatial extent and geometry of major MEQ anomalies, highlighting the densest part of the cluster. The map depicts a dominant MEQ cluster, with its highest density roughly aligning with the southern upflow zone (marked by a star in Figure 2). The most concentrated region, indicated by closely spaced contours, stretches from the Biyong Fault in the southeast to the Tiwi Fault in the northwest, following a predominant northeast-southwest trend. Its longest a*xis* appears to follow the Takla Fault, particularly in its northern segment. Notably, this dense region extends beyond the production area outline, following the southern traces of the Biyong, Takla, and

Tiwi Faults. In contrast, lower-density areas—represented by widely spaced contours—are observed extending just before the Naga Fault to the north and the Bariis Fault to the northwest. Additionally, a separate low-density zone is visible in the southeast, near the smaller blue circle previously shown in Figure 7, which is likely associated with this injection system.

Figure 8B illustrates the density distribution of shallow (represented by contour lines) and deep (depicted by colored contours) MEQs across the Tiwi Geothermal Field. Four distinct dense zones were identified—two associated with deep *events* and two with shallow seismicity. The southern portion of the production area contains the highest concentration of both shallow and deep MEQs, which also coincides with the densest part in Figure 8A. Additionally, northwest of this region, the main production area features a notable cluster of shallow events with minimal deep activity (only contours were displayed). Previous studies have linked deep microseismic events to field injection activities, while shallow events are commonly associated with production operations. Shallow MEQs are interpreted to be triggered by changes in reservoir pressure and subsidence at the surface. Interestingly, the greatest subsidence in Tiwi occurred directly over the main production area. Both shallow and deep MEQs show strong spatial correlation with injection activities in the southeastern part of the field. MEQ swarms frequently follow significant changes in injection rates in this area, reinforcing the connection between injection and seismicity—particularly near the upflow zone and the southeastern injection system. The separate low-density zone identified in Figure 8A is actually characterized by a small, isolated cluster of exclusively deep seismic events, with no notable shallow seismicity.

3.3 OPTICS Parameter Exploration

The OPTICS algorithm was applied to the data set, varying the *min_pts* and *xi* parameters to determine its sensitivity to the clustering solutions and obtain the optimal range of parameters to be used for the analysis.

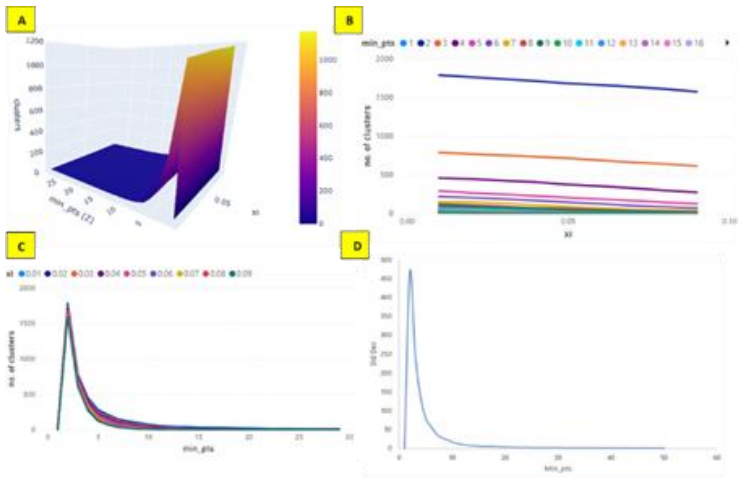

Figure 9: Number of clusters formed by varying *min_pts* and *xi* values

Figure 9A shows a plot of the effect of varying both *min_pts* and *xi* to the number of clusters formed. The number of clusters generally decrease with increasing value of *xi* (see Figure 9B). At smaller values of *xi*, even subtle drops or fluctuations in the reachability of the data are detected which result to fine-grained clusters. At higher *xi* values (approaching 0.1), only the densest clusters are detected. Major clusters can be obtained by using *xi* values 0.02 to 0.08.

On the other hand, an increase in *min_pts* resulted to a decrease in the number of clusters (see Figure 9C). This is expected as *min_pts* dictates the core distances of the data points, consequently affecting the structure of the reachability plot. For very low *min_pts* (1-3), there is an extremely high cluster count suggesting over-segmentation. As *min_pts* is increased to 15, there is a sharp drop in the number of clusters. Beyond *min_pts* >15, the curve started to flatten, suggesting that the dataset only supports a few stable dense regions, ignoring small structures. Beyond *min_pts* = 20, the number of clusters stabilized to less than 10 clusters.

To analyze the stability of the clusters, the standard deviation of the number of clusters over *xi* values for each *min_pts* value were calculated. The stability curve (see Figure 9D) shows very high standard deviation at low *min_pts* values from 1-5, which implies the high sensitivity of the clustering structure to *xi*. Standard deviation sharply decreased, implying increase in stability, from *min_pts* = 5 to 20, while eventually levelling off starting at *min_pts* = 20, showing stable clusters.

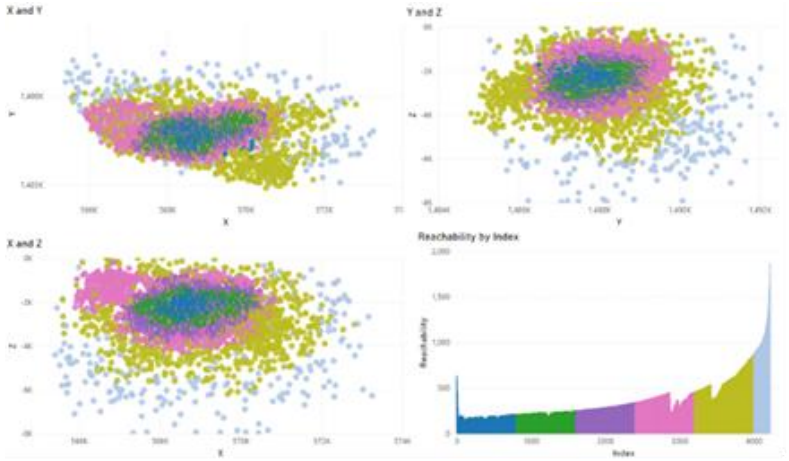

Figure 10: Ordering of points using OPTICS and the corresponding reachability plot (*min_pts* = 15)

The OPTICS algorithm prioritized processing higher density data points, quantified by lower reachability and lower ordering indices (see Figure 10). OPTICS started clustering the data set from the inside (dark blue) then going outwards (light blue), which reveals that data points near the core of the spatial data set have higher density compared to the outer.

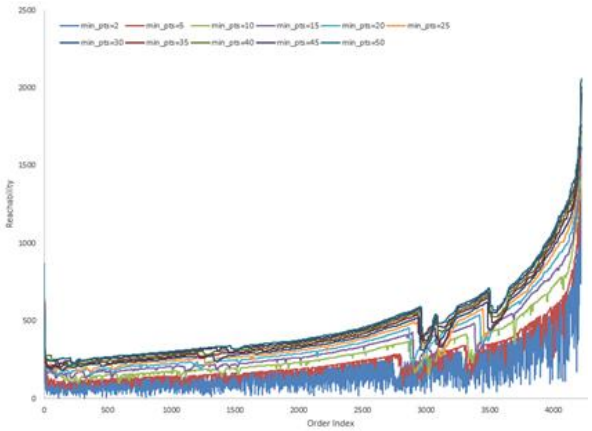

Figure 11: Reachability plot at different *min_pts* values

Figure 11 shows the effect of varying *min_pts* to the reachability plot. Higher reachability values were observed as *min_pts* ¬was increased. As the minimum points required to form a cluster is increased, the core distance is consequently increased.

The valleys are more fine-grained as *min_pts* was decreased below *min_pts*=5, as the criteria for the clustering structure becomes less strict with respect to the number of points. Increasing *min_pts* revealed the main, stable clustering structure of the dataset.

Another feature of the reachability plots generated is the location of the prominent valleys formed, albeit a slight shift in the order by which the steep-downward area started to occur. The delay in the valleys is another consequence of the increased core distances as the *min_pts* is increased. The higher reachability distance delays the 'discovery' of the dense regions in the process of ordering, as more points need to be processed before reaching core points with sufficient density that would result to a valley in the reachability plot.

As the *min_pts* parameter was increased, the valleys formed for order indices less than 1,500 disappear, while the valleys at order indices 2,900-3,300 and 3,500-3,700 remain. The delays in the onset of the steep-downward regions became less prominent as *min_pts* was increased from 20. This indicates that the ordering solution is more stable at *min_pts* > 20.

3.4 OPTICS Main Clustering Structure

Stable 'valleys' which remained in the reachability plot even after varying the *min_pts* parameter was identified by visual inspection of the plot. By visual inspection, the 11 clusters which remain stable even after changing the *min_pts* variable from *min_pts* =20-50 are labeled from A-K in the reachability plot shown in Figure 12. The XYZ coordinates of the labelled valleys were identified and its locations in the Tiwi geothermal area were also mapped out.

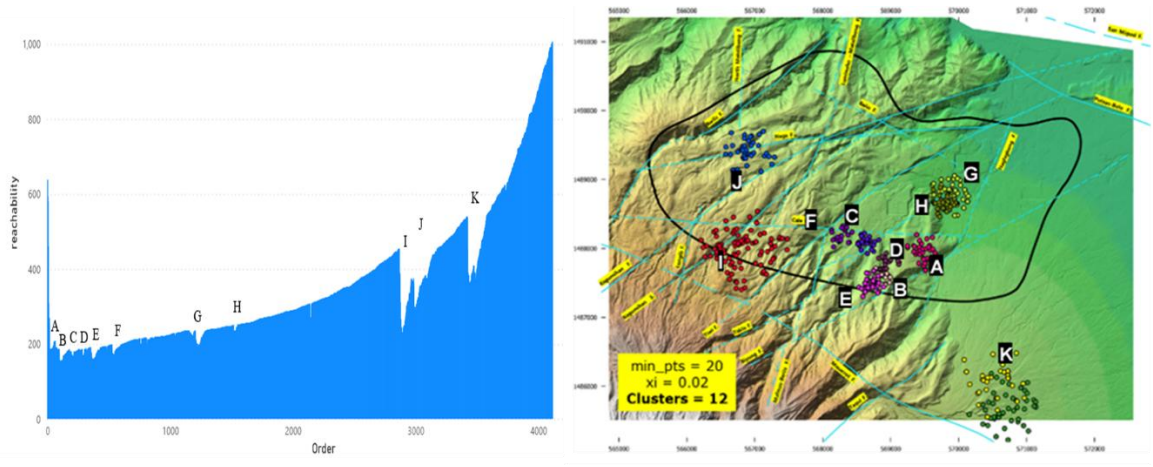

Figure 12: OPTICS Clustering Solution and Reachability plot at *min_pts* = 20

Clusters A–F appear in the low-order index range (<1000) of the reachability plot, indicating high-density and relatively compact structures.

Cluster G was detected only at *min_pts* = 20 and 30. It was relatively large (49 points) at *min_pts* = 20 and *xi* = 0.02–0.03, but reduced in size at higher *min_pts*, dissolving at steeper *xi*, indicating low stability. Cluster H was the least stable, appearing only at *min_pts* = 20 and *xi* = 0.02.

In contrast, Cluster I was the most stable in the dataset, consistently detected across all tested parameter combinations (*min_pts* = 20–50), reflecting a robust and persistent structure. Cluster J appeared for *min_pts* = 20–35 at *xi* = 0.02–0.03. Cluster K, the largest overall, peaked at 149 points at *min_pts* = 35. At lower *min_pts*, it appeared as two separate clusters—one deeper than the other—which merged into a single structure as *min_pts* increased.

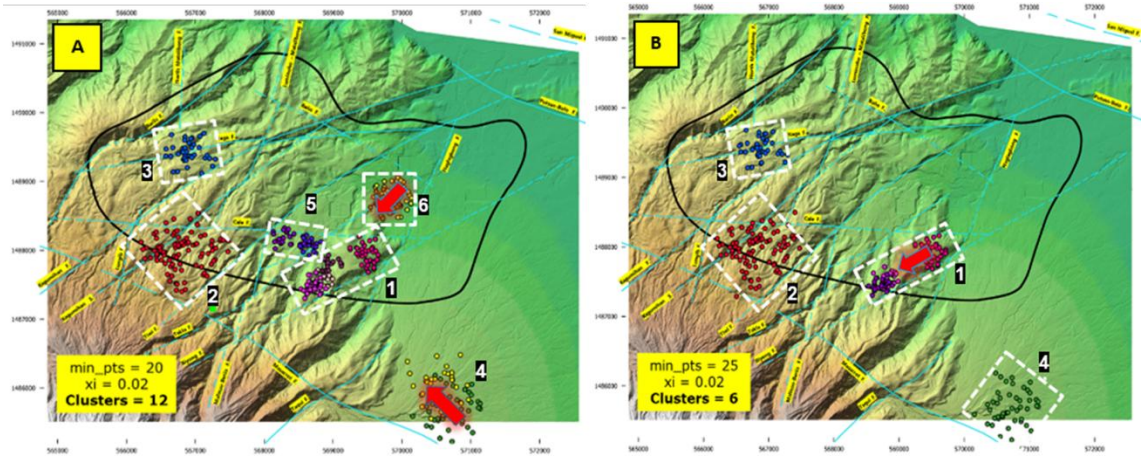

Figure 13: Clusters identified using the OPTICS algorithm. A) Twelve (12) clusters generated with parameters *min_pts* = 20 and *xi* = 0.02; B) Six (6) clusters generated using *min_pts* = 25 and *xi* = 0.02. Arrows show possible injection flowpaths.

The clusters were further grouped into six (6), each identified as separate clusters or sets of clusters, potentially induced by different mechanisms (see Figure 13A). Most of these groupings exhibit some alignment with fault planes—either following their general orientation or clusters adjacent to them—suggesting that faults may serve as structural barriers. These fault zones, being more fractured and mechanically weaker, are more sensitive to cumulative stress changes driven by variations in temperature and pressure.

Clusters with dispersed MEQs are possibly caused by reservoir cooling from marginal recharge. This recharge enters gradually and at a relatively steady and slow rate, causing minimal and dispersed disturbance to the stress field. Group 2 exemplifies this, with shallow MEQs at depths averaging around 900 mbsl, which is within relative entry depths of marginal recharge. The relatively lower pressure west of the Tiwi Fault may have facilitated the inflow of peripheral fluids from surrounding aquifers, driven by the natural movement from areas of higher pressure to lower pressure. This group is approximately centered along the Kagumihan Splay, with the Tiwi Fault forming its southeastern boundary and the Kagumihan Fault marking its northwestern limit. Another mechanism contributing to dispersed MEQ patterns may be the prolonged injection of cooler fluids. This process suggests ongoing changes in the local stress field due to continued cooling of the surrounding rock, as seen in Group 4. This group consists of two clusters: the dark green cluster, located at

greater depths, and the yellow cluster, which is approximately 750 meters shallower. A potential flow path can be inferred, beginning with fluid injection at the deeper dark green cluster, which averages around 3200 mbsl, moving northwestward and upward toward the shallower yellow cluster at approximately 2450 mbsl. This depth gradient suggests a shallowing migration of cooler fluids, potentially influenced by a lithologic or structural barrier that restricts deeper movement toward the northwest. These finding merits further investigation.

In contrast to the dispersed seismic patterns elsewhere in the field, the southern region—highlighted by a star in Figure 2—shows persistent seismic activity marked by tightly clustered events. This pattern is more likely attributed to reservoir-induced processes rather than natural upflow signature of seismicity. These includes Groups 1 and 5, which are likely induced by cooling from deep injections returns, potentially combined with marginal recharge. These lumps of closely spaced events appear to be influenced from high-capacity injection wells located in the southeastern part of the field (see blue circles in Figure 7), where injection depths range from approximately 1900 to 2400 mbsl. The primary distinction between the two groups lies in their fault associations: Group 1 shows more affinity with the Biyong Fault, which bounds it to the southeast, while Group 5 is associated with both the Cale Fault and the Tiwi Fault, forming its northern and northwestern boundaries.

Group 6 consists of two distinct clusters. The light green cluster, located farther from the upflow region than the olive green cluster, lies approximately 650 meters shallower. Injection appears to begin at shallower depths, averaging around 1800 mbsl near the light green cluster, and then progresses southwestward, deepening toward the olive-green cluster at approximately 2450 mbsl. This deepening trend towards the southwest may suggest a possible injection flow path direction towards to upflow zone in the southern part of the field, consistent with pressure contours. Meanwhile, Group 3 possibly reflects a combination of dominant stress changes due to production activities and cooling from injection—possibly originating from the red circles shown in Figure 7—and marginal recharge from peripheral aquifers outside the main production area. This group is generally aligned along the Naga Fault, bounded by the Kagumihan Fault to the southeast and the Bariis Fault to the northwest, with average depths around 1300 mbsl. This area might highlight zones with distributed fracture permeability where stress changes associated to production-related pressure and temperature change as well as subsidence contribute to inducing MEQs

In Figure 13B, four (4) groupings were identified, with Groups 5 and 6 no longer present. The remaining four groupings largely retain the same characteristics as those in Figure 13A, including overall shape and spatial boundaries. Adjusting the parameters in this case only reinforces the presence of the distinct groupings 1, 2, 4, and 5.

Across all parameter variations, the general location, geometry, and bounding faults of these groupings remained consistent, reinforcing the reliability and robustness of the identified seismic patterns.

4 SUMMARY AND CONCLUSION

The OPTICS algorithm was able to detect clusters from the dataset which cannot be delineated by the traditional grid-based technique of clustering MEQs. By parameter exploration, the optimal range of the two input parameters, *min_pts* and *xi* were determined. The OPTICS algorithm successfully identified nested clusters within larger, denser groupings. This adjustment enabled the identification of potential injection flow path directions. The algorithm was also able to highlight the most stable and distinct clusters, satisfying clustering criteria while maintaining spatial constraint by fault zones. Most of these clusters exhibit alignment with fault planes—either by following their orientation or forming adjacent to them—implying that faults may act as structural barriers influencing cluster formation. This spatial limitation suggests that bounding faults may play a significant role in the formation of these clusters. Moreover, OPTICS revealed that even individual clusters tend to follow the general orientation of permeable structures in Tiwi, indicating a strong affinity with fault zones. Through six distinct groupings, the clustering analysis inferred various potential mechanisms of induced seismicity across the field. These insights are difficult to extract using grid-based clustering methods.

The presence of deep clusters in Groups 1 and 5, associated with significant temperature contrasts, suggests not only localized stress changes but also the potential for reactivation of pre-existing faults. The alignment of these clusters with major fault structures (e.g., Biyong and Cale Faults) implies that cooling-induced contraction may be enhancing slip along these faults, contributing to microseismicity.

The occurrence of Group 6 near the northern tip of the Takla Fault, which potentially resembles a horsetail splay, highlights the structural complexity at fault terminations. Horsetail splays are known to act as efficient fluid conduits, and their presence here may facilitate the downward migration of injected fluids toward the upflow zone. This could also explain the relatively shallow depth of Group 6 and its directional flow pattern.

The inferred flow paths from deeper to shallower clusters (e.g., in Groups 4 and 6) may represent thermal front propagation through the reservoir. As cooler fluids move through permeable zones, they alter the thermal gradient, which in turn affects the stress field and triggers MEQs. Monitoring these patterns over time could provide insights into the efficiency and reach of injection strategies.

The consistent spatial boundaries and fault associations across varying parameters suggest that structural controls—rather than random fluid dispersion—govern the distribution of seismic clusters. This reinforces the importance of integrating fault geometry and orientation into reservoir management.

The identification of stable clusters and directional flow paths implies that certain zones within the reservoir are more responsive to injection and recharge. These areas may serve as key targets for sustainable reservoir development, balancing fluid extraction with controlled reinjection to minimize stress perturbations.

5 ACKNOWLEDGEMENTS

The authors would like to acknowledge Philippine Geothermal Production Company, Inc. (PGPC) for the opportunity to share this study with the international geothermal community, contributing to the advancement of knowledge in support of sustainability.

REFERENCES

Agrawal, K. P., Garg, S., Sharma, S., & Patel, P. (2016). Development and validation of OPTICS based spatio-temporal clustering technique. Information Sciences, 369, 388–401. https://doi.org/10.1016/j.ins.2016.06.048

Ankerst, M., Breunig, M. M., Kriegel, H.-P., & Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, 49–60. https://doi.org/10.1145/304182.304187

Calibugan, A., Villaseñor, L., Menzies, A., and Stark, M. (2015). An Evaluation of the Impact of MatRidge Injection to Matalibong Production, Tiwi Geothermal Field, Philippines, Proceedings, World Geothermal Congress.

Ejera, V., Aurelio, M., and Taguibao, K. (2024). 3D modeling and structural setting reconstruction of the Tiwi Geothermal Field using integrated surface and downhole data. GeoCon 2024, Quezon City, Philippines.

Marcuap H., and Villamer, V. (2019). Tiwi Microearthquake (MEQ) Report May 2012 – June 2017. Philippine Geothermal Production Company Internal Report.

Marcuap, H., and Villamer, V. (2017). Tiwi Microearthquake (MEQ) 2016 Annual Report. Philippine Geothermal Production Company, Inc. Internal Report.

Mossop, A., and Segall, P. (1997). Induced seismicity in geothermal fields: Thermoelastic Injection model. Journal of Geophysical Research.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830. http://scikit-learn.sourceforge.net

Piegari, E., Herrmann, M., & Marzocchi, W. (2022). 3-D spatial cluster analysis of seismic sequences through density-based algorithms. Geophysical Journal International, 230(3), 2073–2088. https://doi.org/10.1093/gji/ggac160

Schubert, E., & Gertz, M. (2018). Improving the cluster structure extracted from OPTICS plots. Proceedings of the 20th International Conference on Extending Database Technology (EDBT 2018), 341–352. https://doi.org/10.5441/002/edbt.2018.30

V, N. C., & Surendran, M. (n.d.). Review of spatial clustering methods. International Journal of Information Technology Infrastructure, 2(3). http://warse.org/pdfs/2013/ijiti01232013.pdf