Relationships between tectonic environments and deposits of gold in Western U.S.A displayed through Random Forest and Support Vector Machine

J. Chou, A. Nasir, A. Dahal

Abstract

Gold deposits that occur along the western coast of U.S.A are unique in their depositional setting due to the variance in tectonic activities that have occurred over time. The importance of subduction zones in deposits is simple. As converging plates collide along their boundaries in a subduction zone, energy gets released along them to produce magmatism and create volcanoes and earthquakes. These create the ideal environment for minerals to precipitate and concentrate, resulting in large formations of porphyry deposits. For our topic, we analyse the western side of North America to determine the relationship between tectonic processes and gold deposits. By using data of known gold deposits from the data repository EarthChem, we utilise the programming language Python to gain geospatial and temporal information such as mineral abundance in relation to the age of deposits and location of ore formation. Further, by applying machine learning techniques such as random forest (RF) and support vector machine (SVM) we compare information on a set of variables to gain optimised data on the types of tectonic processes that occur and how they impact on our deposits. With the machine learning models, we can create visual spatial data for subduction-related episodes such as convergence rates of the plates. This ultimately assists us in deriving and understanding the causal relationship between certain parameters and our deposit data.

1. Introduction

Gold is an invaluable resource and has a plethora of industrial, economic and electronic applications due to its high malleability, conductivity, resistance to corrosion, and investment status as a long term store of value. Gold is a relatively scarce element, with economically viable deposits of gold forming as either lode (or primary) deposits, and placer (or secondary) deposits (Kirkemo, Newman and Ashley, 2019). Deposits of gold are distributed all over the world, with the largest economic reserves located in Australia, South Africa and Russia (Geoscience Australia, 2013). As of 2013, the United States of America (U.S.A.) was the third largest producer of gold, accounting for 230 tonnes or 8% of total world production (Geoscience Australia, 2013).

The tectonic environment of western U.S.A. and its relationship to the formation of gold deposits will be explored through two different machine learning methods. Our obtained gold deposit dataset was geographically constrained to focus on the western region of North America. The western side of the U.S.A. was chosen due to its unique tectonic setting. These settings are a result of tectonic processes and impact the geological environment in interesting ways, creating depositional environments for economically viable minerals. For this project we firstly examined the North American Cordillera, focusing on the western coast near California. By examining the Cordillera we can explore the development of the retro arc basin system and the foreland basin system of western North America and its impact on the formation of gold deposits. The Western Cordillera of North America is a major system of mountain ranges and plateaus, extending through the U.S.A, from Canada to Mexico. The deposit types near the cordillera are dominated by porphyries (Singer et. al., 2008), with sediment-hosted and high-sulfidation epithermal deposits also present (Sillitoe, 2008). It is estimated that the gold belts and deposits found in the Western Cordillera of North America contain over 10 Moz or 300 tonnes of economically viable gold (Sillitoe, 2008). The majority of deposits that formed at this location are a result of extensional tectonic and magmatic conditions, near arc or back-arc basins (Sillitoe, 2008). The deposits predominantly occur along the edge of the North American craton and were formed after the Cordilleran orogenic belt became tectonically consolidated after the accretion of fringing arcs during the Late Jurassic approximately 155 Ma (DeCelles, 2004; Sillitoe, 2008). After 30Ma, a transform boundary was present near the coast of California and a strike-slip thrust fault occurred (Garfunkel, 1973). This type of fault is defined as a right-lateral strike slip fault which means that if you face along the other side of the fault, it appears as though the ground has shifted to the right past the fault line. The thrust fault, known as the San Andreas Fault, prevented other tectonic processes of oceanic plate subduction from occurring as the transform boundaries are only slipping past another, rather than converging and subducting. The level of hydrothermal activity was reduced, compared to the higher levels that are attributed to subducting plates. Consequently the formation of porphyry gold deposits past the fault was significantly restricted and occurred in small volumes (Murray, Torvela and Bills, 2019).

Faults such as the San Andreas fault have an impact on the deposits of ores nearby (Greenbank and Knight, 2013). By understanding the tectonic processes that formed the Western Cordillera of North America and the later event that caused the San Andreas fault, the depositional history and formation of gold belts and deposits in western U.S.A can be understood. From here we utilise information about gold deposits across the world, sourced from EarthByte.com to create figures, charts and tables for temporal and spatial analysis. Specifically, we focus on the area of western U.S.A. and utilise data on variables such as convergence rates, age of the deposits and seafloor age. The information is also cross-examined with data from the GPlates Portal (http://portal.gplates.org/cesium/?view=AgeGrid) (Müller et. al., 2016) where we can examine subduction events that occurred over time. For this specific topic we make use of the age grid for optimal results on subduction zones in relation to age, focusing on the North American region. Through the use of machine learning techniques and the programming language Python, we compare tectonic processes, examine causal relationships and produce spatial and temporal plots of gold deposits. The machine learning processes utilized for this particular project were random forest and support vector machine (SVM) methods. Through this we aim to show new mechanisms of extracting geological and geophysical processes and further our understanding of the relationship between tectonic processes and the formation of gold deposits in Western U.S.A.

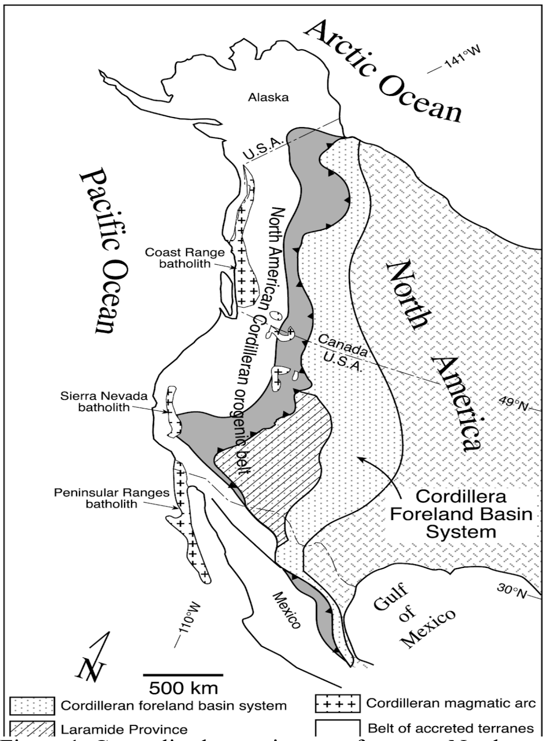


Figure 1. Generalised tectonic map of western North America, showing the major geologic zones of the Cordilleran orogenic belt and foreland basin system (DeCelles, 2004).

1.1. Brief Tectonic Background of Western U.S.A.

The western coast of the U.S.A. has a rich and dynamic tectonic history. The Orogenic belt reaches the length of 6,000 km that extends from Mexico to Canada (Pfiffner and Gonzalez, 2013). Around the Late Jurassic period (155 Ma), the orogenic belt of the Cordilleran system was tectonically consolidated due to the closure of ocean basins and accretion of fringing arcs. This was followed by the large scale contractile deformation of the foreland basin system 1000 kilometres eastward into the North American continent until approximately 55 Ma, resulting in the formation of the Laramide Rocky Mountain Ranges as seen in Figure 1 (DeCelles, 2004). The formation of this retroarc thrust and mountain belt has been used as a typical example of subduction related mountain forming processes. When the Pacific and North American plates converged, the oceanic Pacific plate was subducted underneath the continental North American plate (Dickinson, 2004). The Cordilleran retro arc thrust belt and foreland basin system formed on the continental North American plate during collision (DeCelles, 2004). This temporally and spatially large scale tectonic event was vital in the formation and mineralisation of gold deposits in this region due to the partial melting and magmatism associated with subduction zones (Sillitoe, 2008).

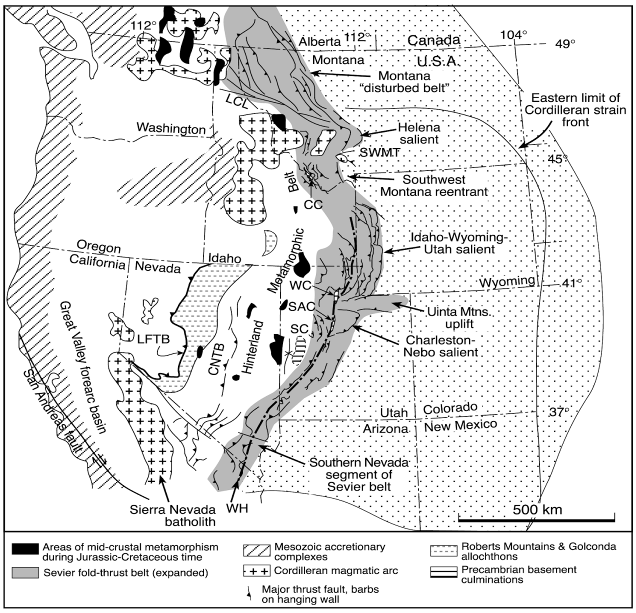


Figure 2. Tectonic map of the western United States, showing the major components of the Cordilleran orogenic belt (DeCelles, 2004).

This was then followed by a secondary tectonic event that occurred during the mid-Cenozoic period (30 Ma), illustrated in Figure 2. When the Pacific and North American plate started to converge, the previous plate named the 'Farallon plate' started to become subducted underneath the North American plate (Garfunkel, 1973). The relative motion of the plate began to shift and rather than converging and subducting, a transform boundary was formed, which is now known as the 'San Andreas Fault'. The new formed transform boundary (also known as strike-slip fault) made a vast impact on the formation of porphyry deposits near the area (Murray, Torvela and Bills, 2019). Gold deposits around the area began to be smaller and much more localised than previous deposits. This is as there is smaller hydrothermal activity near strike slip faults than converging or extensional faults. Following this, the environment needed for minerals and deposits to form become less suited than other areas where there is subduction, high levels of hydrothermal activity and magmatism. The North American Cordillera that resulted in larger scale, concentrated gold deposits was followed by a transform boundary which changed the area to have more localised deposits.

2. Data Set for Gold Deposits from EarthChem

The initial data set for gold deposits along the western coast of U.S.A. was imported through a database called EarthChem (https://www.earthchem.org/) (Lamont-Doherty Earth Observatory of Columbia University, 2019). The majority of the data provided for gold from EarthChem seems to be derived from USGS and NAVDAT databases (Lamont-Doherty Earth Observatory of Columbia University, 2019). Variables such as the age of deposits, abundance, latitude and longitude are given for gold deposits around the globe. The data had to be synthesised to locations around present day North American continent. With these gold deposits we see that they are mostly situated along 50°N to 30°N. The relationship between geospatial and geo temporal abundance of these deposits are the unique factor to be analysed in this paper. Figure 3 illustrates the deposits of gold and its geospatial location in relation to the age of deposits and is an example of the plots produced. The deposits are coloured to show their age and the distance along the trench.

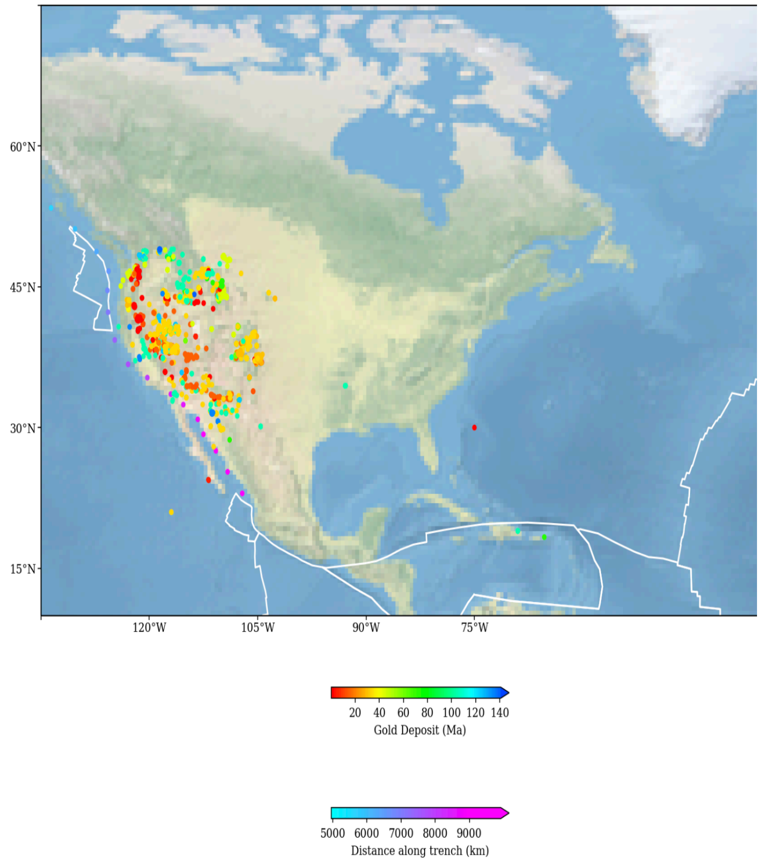


Figure 3. Geospatial plot created using Python illustrating location and age of gold deposits in western North America.

3. Python Methods

The programming language Python was used to produce the vital figures, tables and charts needed to visualise the tectonic processes that occurred in our area of focus. Firstly, we created environments and obtained kinematic data. We began by creating environments for Conda which allows for the python environments needed for the script. These are environments and modules named matplotlib, scikit-learn, scipy, python, numpy, jupyter, cartopy, pandas and notebook. To run the kinematic data, we focused the data for ages up to 230 Ma, as to restrict our data set for optimal running.

We chose a restriction of 230 Ma as we hypothesized formation of gold deposits would coincide with the tectonic processes in the formation of the North American Cordillera and the San Andreas Fault, which began approximately 155 Ma (DeCelles, 2004). Additionally, from the histogram above (Figure 4) produced as a result of statistical analysis of our dataset, there is a significant increase in the abundance of gold deposits after approximately 250 Ma. By

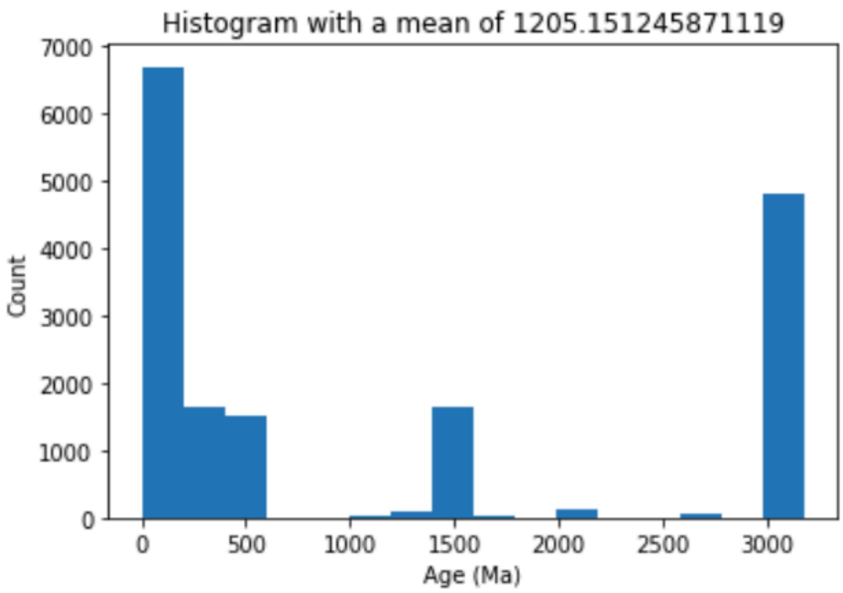


Figure 4. Histogram created using Python illustrating the distribution of the abundance and age of gold deposits in western North America. Our dataset includes gold deposits of age 3000 Ma which is irrelevant for our project.

using Muller's (et al. 2019) convergence file, we are able to extract subduction values for our deposits for the restricted ages. Additionally, the new stats were co-registered with the EarthChem deposit data of gold. A variety of Python functions that were utilised to combine our EarthChem data with the kinematic subduction information provided by Butterworth (et al. 2019). We manipulated the *coregLoop.ipynb* to co-register our gold data taken from EarthChem to be in a readable format compatible with the script and do all the co-registration by assigning given and random ages. Moreover, we manually created a list of arrays containing the Plate ID for the North America Craton since it was missing from the gold data set. After the process of co-registration, a new format of files was created, named pickle, which helps with machine learning. We utilised Butterworth and Chandra (2019) python scripts to co-register points between the shapefile and our text file of EarthChem data which was then imported into Butterworth's (et al. 2019) python notebook which allows us to make use of two machine learning processes. In the machine learning python notebook, we made use of the machine learning codes to apply our porphyry deposit in the western part of North America. Through synthesising our data with the kinematic data, we were able to obtain over 20 different variables we could test to carry out geospatial analysis on our chosen area and deposit. The data set first had to be manipulated into a format that the machine learning code could read so that the scikit.learn environment could be utilised properly. The data was partitioned into training and testing points for training purpose using the classifiers to determine which tectonic parameters are linked to deposits and non-deposits (Butterworth et. al., 2016). The purpose of the "training" set is to learn which tectonic parameters are essential for gold deposits, and later its validity tested with the "testing" set (Butterworth and Chandra, 2019). The machine learning methods used for the purpose of this paper were random forest and support vector machine (SVM). By allowing us to create parameters with which we can focus on specific variables, the primary task was to locate variables that best suited our needs.

3.1. Random Forest (RF) and Support Vector Machine (SVM)

Random forest is a method of classification that combines various decision trees that trains and outputs parameters. By combining these decision trees it can create a final prediction that averages each of the decision trees. We utilised this particular machine learning method as it works best for a cluster of information that is similar to ours. This type of machine learning has been used previously by geological teams (Cracknell and Reading, 2013).

Similarly, support vector machine also works with non-linear data, where it utilises support vectors to separate data in a linear manner. This is done by dividing cases into positive and negative cases (Cortes and Vapnik, 1995). SVM uses kernels to transform the given data into possible outputs. For this dataset, SVM can be considered a favourable and compatible method as it is designed to help with classification of data.

4. Data Results

4.1. Cross-Examination of Age Related Deposits with GPlates

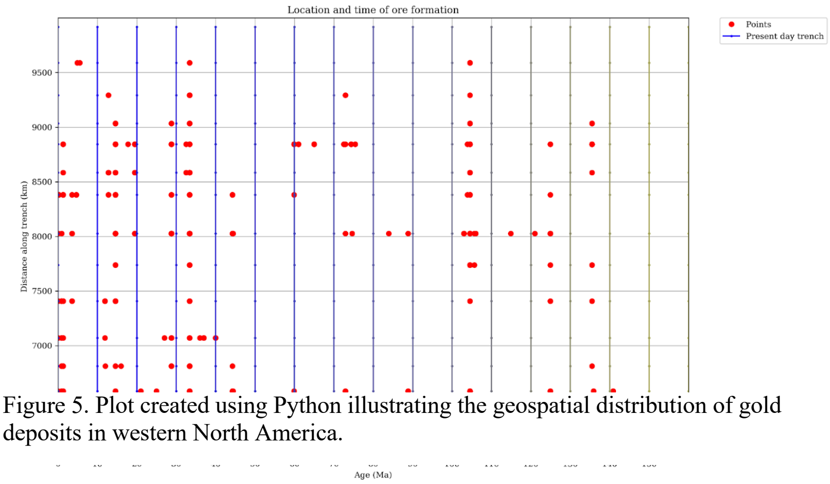


Figure 5. Plot created using Python illustrating the geospatial distribution of gold deposits in western North America.

For our first point of data examination, we obtain the developed map of North America and table of ore formation produced by Python and compared it to the information observed through GPlates simulation. Figure 3, presents the geospatial placement of gold deposits with colour defining its geotemporal relationship. Figures 6, 7 and 8 are screenshots taken from the GPlates portal and display the formation of the fault and the subduction zone over the last 35 Ma.

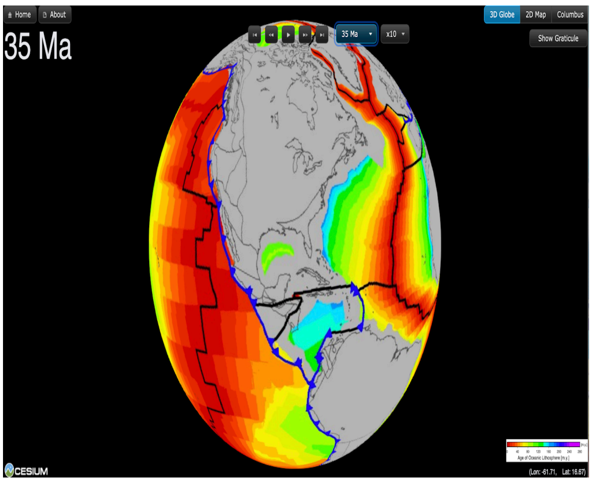


Figure 6. Tectonic plate movement in North America and age of plates at 35 Ma, retrieved from GPlates Age Grid Portal (Müller et. al., 2016)

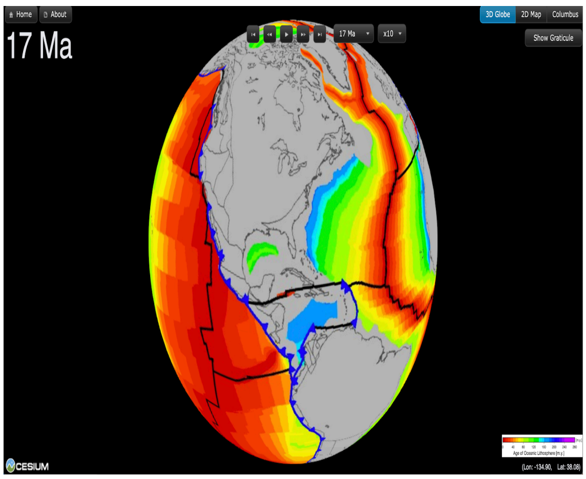


Figure 7. Tectonic plate movement in North America and age of plates at 17 Ma, retrieved from GPlates Age Grid Portal (Müller et. al., 2016)

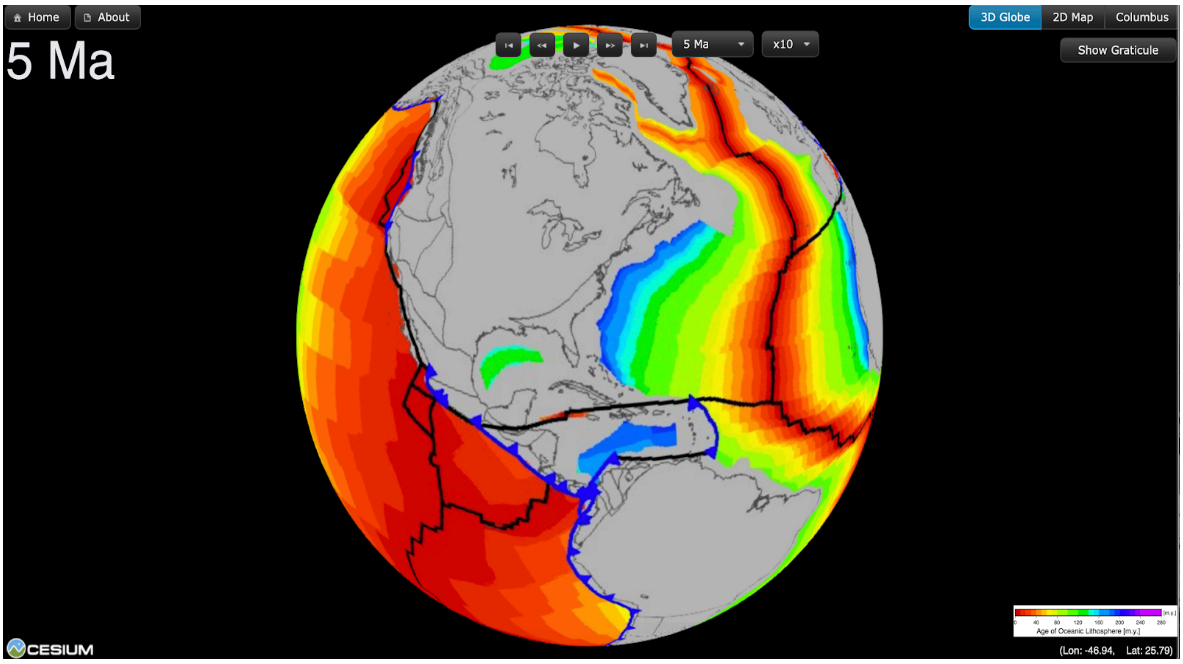


Figure 8. Tectonic plate movement in North America and age of plates at 5 Ma, retrieved from GPlates Age Grid Portal (Müller et. al., 2016)

4.2. Optimal Parameters for Random Forest

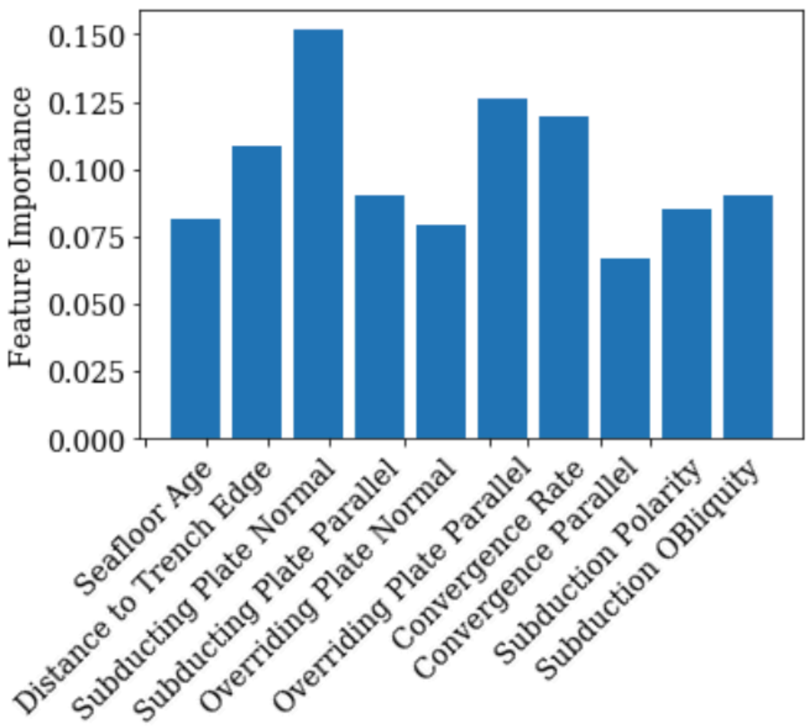


Figure 9. Graph created using Python of parameters present in dataset and the importance of each parameter or feature.

There were a total of 20 RF parameters that could be tested and trained in the machine learning system. To prevent over interpretations and create a better analysis of our deposit data, we chose to test each parameter separately. Random forest's ability to work off of decision trees allows it to test and determine which parameters would be most suitable for a given set of data by testing individual parameters and predicting the relationship with the data (Joly, 2017). As shown in Figure 9, the

RF parameters that would work best with our given dataset would be the parameters that have the highest feature importance compared to other parameters. The parameters when tested separately could appear to have fluctuations, but when comparing the datasets created by random forest, they show similarities. Their significance also begins to show after including supporting evidence, and testing against similar parameters that appear to be of the same level of importance. RF determines the importance of parameters by training and testing the information on each parameter against positive and negative values. In this case, positive values relate to deposits and negative values are the non-deposits (Butterworth et. al., 2016).

4.3. Tectonic parameters and gold deposits

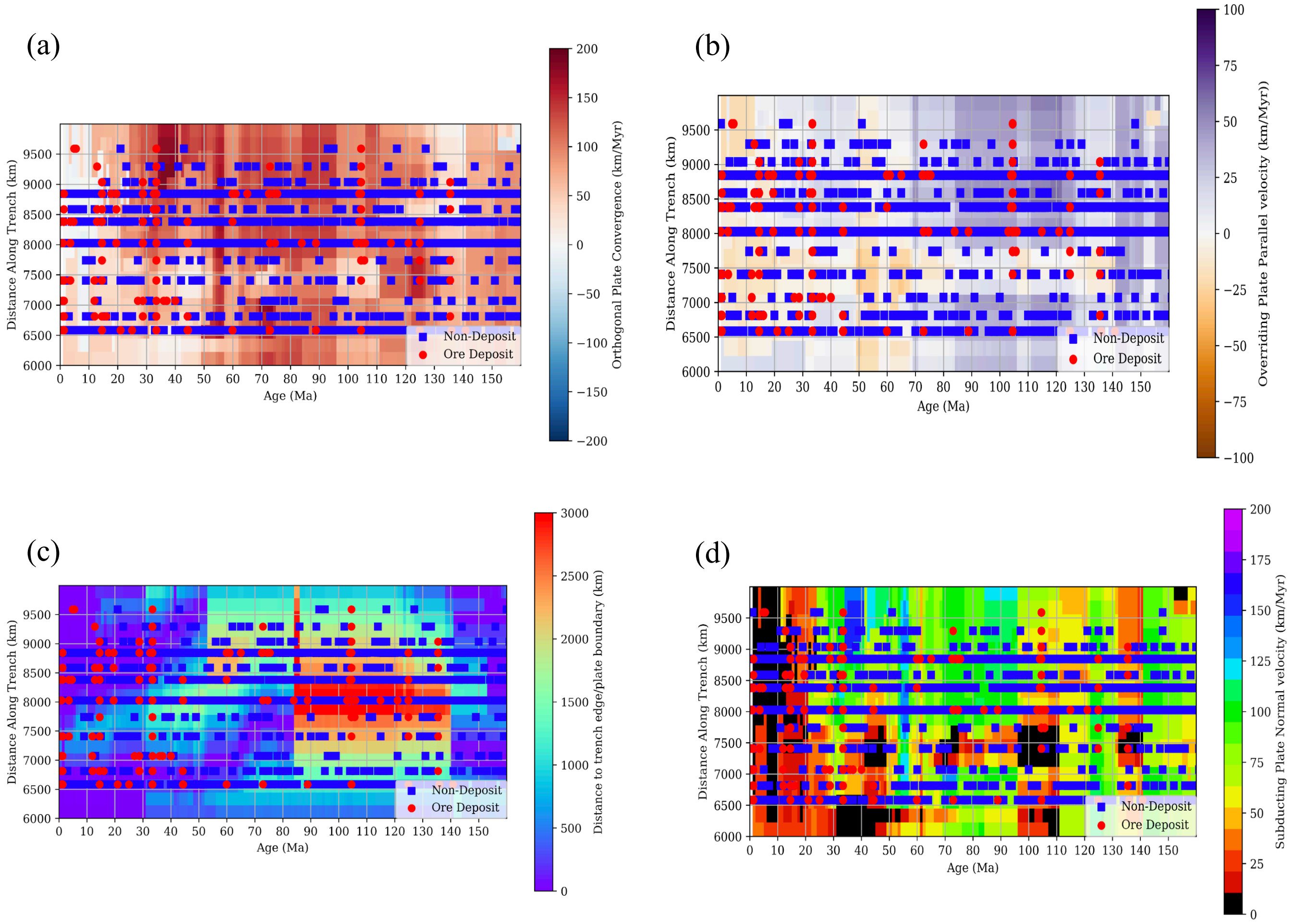
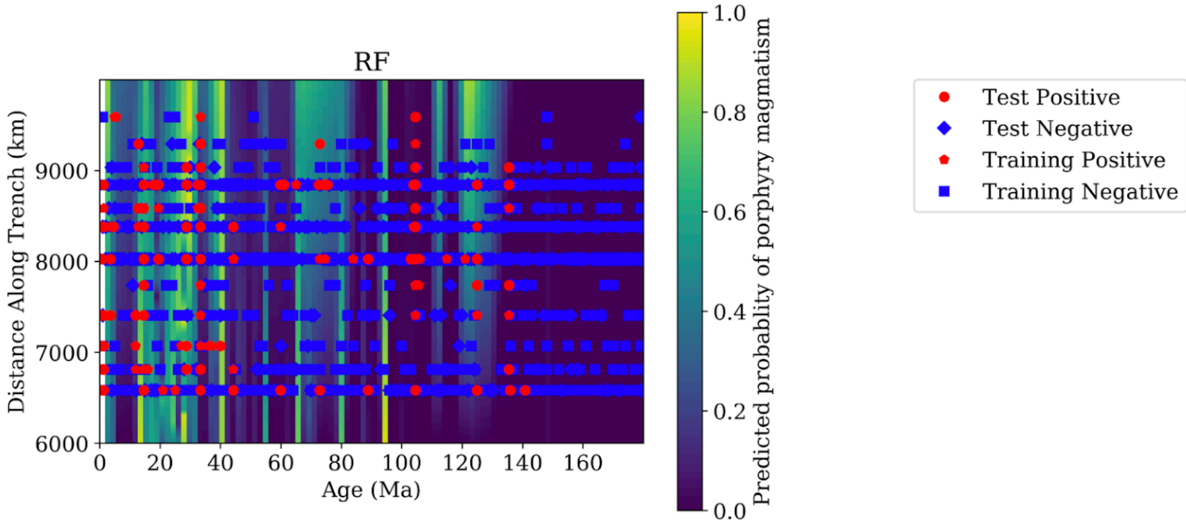


Figure 10. Plots produced as a result of RF machine learning method using Python illustrating relationship between deposit locations and specific parameters. a. Orthogonal Plate Convergence, b. Overriding plate parallel velocity, c. Distance to trench edge/plate boundary and d. Subducting plate normal velocity

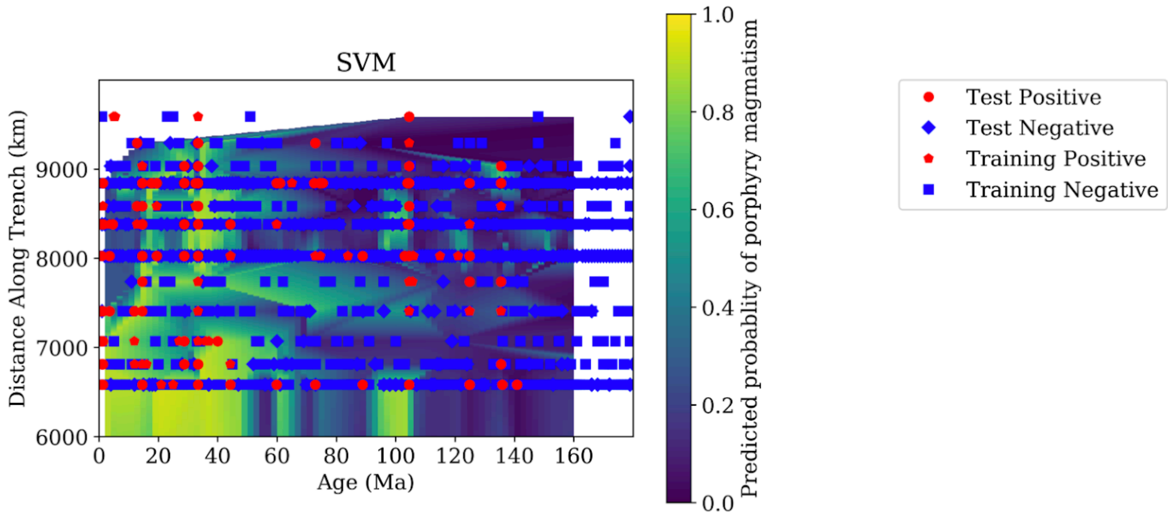
Figure 10 shows the relationship between deposit locations and specific parameters. In these figures, red circles determine our deposit information and the blue circles show non-deposits. Non-deposits are pseudo-random datasets that the machine learning uses to train and test data in a proper algorithm. By providing it with non-deposits, it is able to correctly categorise and find parameters that are negative values (Akbani et. al., 2004).

4.4. Machine Learning Results




```
Five-fold cross validation scores:  [0.87146974 0.90605187 0.89740634 0.96770473 0.89965398]
SCORE Mean: 0.91 STD: 0.03
```

Figure 11. Graph created using Python of results derived from RF machine learning method.




```
Five-fold cross validation scores:  [0.73544669 0.82824207 0.82536023 0.84025375 0.72260669]
SCORE Mean: 0.79 STD: 0.05
```

Figure 12. Graph created using Python of results derived from SVM machine learning method.

Figures 11 and 12 are used to compare the two different machine learning methods. The results are prediction of the actual data. According to Butterworth et. al. (2016), the five-fold cross validation scores are probabilities that represent its accurateness to the actual data. There are five scores because the training and testing sets were split into five sets during data partitioning. Even though both of the data were similar, there is a slight difference in the way SVM and Random Forest predicts the probability of magmatism in our porphyry data. Although we can expect the deposit data to be in the yellow zone with more magmatism, SVM predicts there to be more magmatism closer to present day than RF does. The 'test' and 'train' datasets refers to the way these machine learning programs learn and improve datasets. The testing and training between both of the machine learning programs seems to be similar, with the only difference being that support vector machine seems to go over the boundary layer (possibly hinting on extra testing and training models).

5. Discussion

Greenbank and Knight (2013) state that transform boundaries cause deposits that form nearby to be smaller and localised as opposed to subduction boundaries. One of the gold deposits that may be an example of this is the hot spring gold deposit located in McLaughlin, close to the Californian coast (Sherlock & Lehrman, 1995). This compared to Figure 5 appears to be the case with only a small room of error. The placements of gold deposits in recent years (prior to 30 Ma) appears to be focused near the coast line, localising near the fault, with only 2 or 3 deposits occurring away from the coast. To help visualise, Figures 6, 7 and 8 display the formation of the fault and the removal of the subduction zone. When cross compared to the age map created in python figure 3, we can establish a relationship with the tectonic process and the deposits of gold. Figure 3 displays the location and time of the formation of gold deposits. At the time of the active tectonic events of the Cordillera, there seem to be various formations of gold deposits, but they are further away from current day trench. The data shows that post events that created San Andreas fault, there is an abundance of gold deposits along present day trench. Noticeably, there seems to be a lack of deposits near the trench during 40-100 Ma. This could be due to errors, but when examining in relation to figure 5, we can see that the deposits that occurred during that time frame, were further inland, than closer to the trench.

5.1. Relationship between Tectonic Parameters and Gold Deposits

Figure 10(a) shows ore deposits peaking when convergence rate is near 0 km/Myr for 0-30 Ma. This could be indicating the effects of the San Andreas fault, whereas for when the oceanic plate was still subducting (before the fault formed), the deposits appear on zones that are converging at 50-100 km/myr. 10(b) indicates that during the phase of faulting, the deposits occur at areas where the parallel plates are 0 to -10km/Myr. Past 30Ma, during the subducting phase, the deposits appear to be generally placed around 30 km/Myr. Furthermore, Figure 10(c) depicts deposits that seem to be placed around 0km in between 0 to 30 Ma. This is followed by a change where past 30 Ma, they appear to be over 2000 km. This is possibly due to there being no subduction events that occur from 0-30 Ma, due to the faulting. Lastly, Figure 10(d) depicts similar changes during the 2 different phases. During 0-30 Ma, deposits peak around 10 km/Myr velocity whereas past 30Ma, they seem to peak around 30-60 km/Myr.

The results of the machine learning tables all seem to link in that they do portray different sets of deposit data. The deposit setting after the formation of the San Andreas fault does appear to be different to the processes before the fault. As there is no subduction occurring along the boundary, there is less velocity of plates, and the overriding plates appear in a parallel form.

5.2. Machine Learning

The primary tectonic parameters that appeared to be relevant to our deposit data were distance to trench age, subducting plate normal velocity, overriding plate parallel velocity and Convergence normal rate. These sets were chosen as they appear to have the highest level of importance on gold deposits along the western part of North America. Although, one limitation of these parameters are that gold deposits can also be impacted on by variables that are not considered in this dataset. These can be aspects such as the presence of water, temperature and more. The sets were then tested and trained through random forest and support vector machine to categorise the deposit data into positive and negative values and then to create as accurate data as it could. Due to the unique setting of tectonic history of our location, these parameters helped quantifiably portray data to separate the deposits relationship with its geospatial relationship over time. The results from Figure 10 show how there is difference between 0 to 30 Ma and past 30 Ma. Although, the data appears to have errors in them which could be causing there to be issues in accuracy. The results that came from the machine learning did not seem to show any signs of impact from the cordilleran orogenic system. This could be due to misinterpretation or the nature of the data.

5.3. Limitations of our datasets

The datasets for gold deposits that were taken from Earthchem.org appears to be recorded inexplicitly. This can be observed in Figure 5 in how the distribution of the data is in line like a grid. The data were not recorded to greater accuracy. This had causes the non-deposits to be distributed in a line as well, due to this limitation because non-deposit points depend on deposit as based on Butterworth et. al. (2016), the non-deposit data is created based on the ore deposit current location by allocating random ages of formation. The limitation of age in our datasets (as it only goes up to 150 Ma), could be a limitation in testing against 3 different types of tectonic processes. As there were 2 major tectonic events in our location, having an extended amount of time to measure could have given the edge for more accurate interpretation of data.

6. Conclusion

To conclude, the analysis of geo temporal and geospatial relationship between gold deposits and tectonic processes along western U.S.A. proved to be challenging due to the nature of the provided data. By using random forest, and support vector machine learning we test and train the results as best as we can to categorise the results and to show the physical relationships of the results. The predicted results show a shift in deposit data related to the tectonic event that occurred in the last 30 Million years. The creation of the transform boundary ceased all subduction near the western coast of U.S.A. All the data and analysis seems to show indication of these changes, but due to the nature of the data, there were various limitations of our analysis. The impact of the cordillera system did not seem to have as big of an affect as the San Andreas fault, which seems suspicious and could be due to the nature of data provided.

7. Appendix

For additional information on the codes that we utilised for creating maps, and the machine learning processes, here is our github repository.
https://github.com/Aeshes1/GEOS3888

References

Akbani, R., S. Kwek, and N. Japkowicz (2004). Machine Learning ECML, Chap. Applying Support Vector Machines to Imbalanced Datasets, pp. 39–50, Springer, Berlin.

Butterworth, N., Steinberg, D., Müller, R. D., Williams, S., Merdith, A. S., and Hardy, S. (2016). Tectonic environments of South American porphyry copper magmatism through time revealed by spatiotemporal data mining. *Tectonics*. [Online] 35 (12), 2847–2862.

Butterworth, N. and Chandra, R (2019). GEOS3888: *Economic Geology Interdisciplinary Project, week 10, mineralexploration course data* [code and data]. Retrieved from https://github.com/intelligentexploration/ Mineralexplorationcourse

Cortes, C., and Vapnik V. (1995). Support-vector networks, Mach. Learn., 20(3), 273–297, doi:10.1007/BF00994018

Cracknell, M. J., and Reading A. M. (2013). The upside of uncertainty: Identification of lithology contact zones from airborne geophysics and satellite data using random forests and support vector machines, *Geophysics*, 78(3), WB113–WB126, doi:10.1190/geo2012-0411.1.

DeCelles, P. (2004). Late Jurassic to Eocene evolution of the Cordilleran thrust belt and foreland basin system, western U.S.A. *American Journal Of Science*, *304*(2), 105-168. doi: 10.2475/ajs.304.2.105

Dickinson, W. R. (2004). Evolution of the North American Cordillera: Annual Review of Earth and Planetary Sciences,v. 32, doi:10.1146/annurev.earth.32.101802.120257, p. 13–44.

Garfunkel, Z. (1973). History of the San Andreas Fault as a Plate Boundary, *Bulletin of the Geological Society of America,* vol. 84, no. 6, pp. 2035-2042.

Geoscience Australia (2013). *Gold*. [online] Geoscience Australia. Available at: https://www.ga.gov.au/data-pubs/data-and-publications-search/publications/australian-minerals-resource-assessment/gold [Accessed 29 May 2019].

Greenbank, L. and Knight, A. (2013). *Fault lines lead to gold - CSIROscope*. [online] CSIROscope. Available at: https://blog.csiro.au/fault-lines-lead-to-gold/ [Accessed 14 May 2019].

Jeppson, T.N., Bradbury, K.K. & Evans, J.P. (2010). Geophysical properties within the San Andreas Fault Zone at the San Andreas Fault Observatory at Depth and their relationships to rock properties and fault zone structure, *Journal of Geophysical Research.Solid Earth,* vol. 115, no. 12.

Joly, A. (2017). *Exploiting random projections and sparsity with random forests and gradient boosting methods -- Application to multi-label and multi-output learning, random forest model compression and leveraging input sparsity*, Cornell University Library, arXiv.org, Ithaca.

Kirkemo, H., Newman, W. and Ashley, R. (2019). *Gold*. [online] USGS. Available at: https://pubs.usgs.gov/gip/prospect1/goldgip.html [Accessed 29 May 2019].

Lamont-Doherty Earth Observatory of Columbia University. (2019). EarthChem Portal. Retrieved from http://www.earthchem.org/portal

Müller, R., Seton, M., Zahirovic, S., Williams, S., Matthews, K., Wright, N., Shephard, G., Maloney, K., Barnett-Moore, N., Hosseinpour, M., Bower, D. and Cannon, J. (2016). *Seafloor-Spreading Age Grids of The Ocean Floor*. [online] GPlates Portal. Available at: http://portal.gplates.org/cesium/?view=AgeGrid [Accessed 29 May 2019].

Murray, S., Torvela, T. and Bills, H. (2019). A geostatistical approach to analyzing gold distribution controlled by large-scale fault systems – An example from Côte d'Ivoire. *Journal of African Earth Sciences*, 151, pp.351-370.

Sillitoe, R. (2008). Special Paper: Major Gold Deposits and Belts of the North and South American Cordillera: Distribution, Tectonomagmatic Settings, and Metallogenic Considerations. *Economic Geology*, *103*(4), 663-687. doi: 10.2113/gsecongeo.103.4.663

Sherlock, R.L. & Lehrman, N.J. (1995). 'Occurrences of dendritic gold at the McLaughlin Mine hot-spring gold deposit', *Mineralium Deposita*, vol. 30, no. 3–4, pp. 323–7.

Singer, D., Berger, V. and Moring B. (2008). Porphyry copper deposits of the world: Database and grade and tonnage models, U.S. Geol. Surv. Open File Rep., 2008–1155.

Pfiffner, O.A. & Gonzalez, L. (2013). "Mesozoic-Cenozoic Evolution of the Western Margin of South America: Case Study of the Peruvian Andes", *Geosciences,* vol. 3, no. 2, pp. 262-310.

Wilson, D., McCrory, P. and Stanley, R. (2005). Implications of volcanism in coastal California for the Neogene deformation history of western North America. *Tectonics*, 24(3).