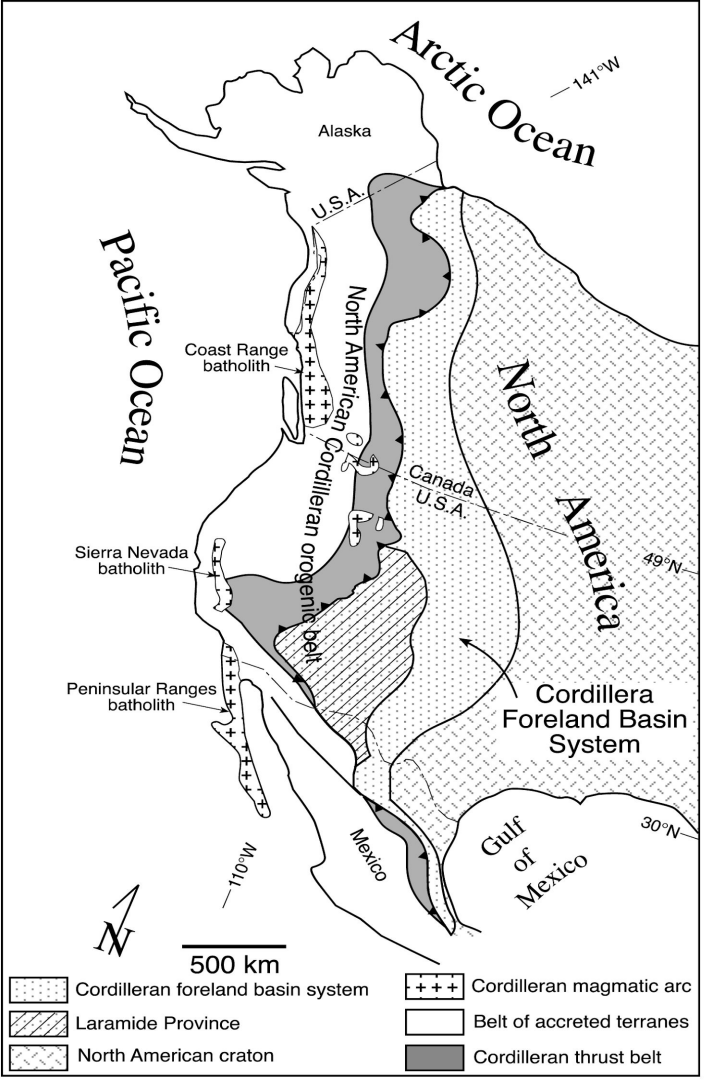
Relationships between tectonic environments and deposits of gold in Western U.S.A displayed through Random Forest and SVM

Jacinda Chou, Amir khwan Haniff Mohd Nasir and Ashesh Dahal


NORTH AMERICA
ASIA
EUROPE
SOUTH AMERICA
ARCTIC OCEAN
PACIFIC OCEAN
ATLANTIC OCEAN
Caribbean Sea
Greenland (Kalaallit Nunaat) (Denmark)
Greenland Sea
Baffin Bay
Davis Strait
Denmark Strait
Labrador Sea
Hudson Bay
Hudson Strait
Foxe Basin
Gulf of Boothia
Baffin Island
Queen Elizabeth Islands
Ellesmere Island
Nares Strait
Arctic Circle
Chukchi Sea
Bering Strait
Bering sea
Beaufort Sea
Alaska (US)
Brooks Range
Alaska Range
Mount McKinley 6194 m
Aleutian Range
Yukon
Mackenzie Mts.
Mackenzie
Banks Island
Victoria Island
Great Bear Lake
Great Slave Lake
Lake Athabasca
Reindeer Lake
Gulf of Alaska
Alexander Archipelago
Yukon Coast Mts.
Rocky Mountains
CANADA
Edmonton
Lake Manitoba
Lake Winnipeg
Winnipeg
Calgary
Vancouver
Seattle
Portland
Mt. Baker 3286 m
Queen Charlotte Islands
Mt. Hood 3426 m
Grand Teton 4199 m
Great Plains
Missouri
Mississippi
Lake Superior
Lake Nipigon
Québec City
Montréal
OTTAWA
Toronto
Lake Ontario
Lake Erie
Detroit
Chicago
Indianapolis
Ohio
Boston
New York
Philadelphia
WASHINGTON, D.C.
Mt. Katahdin 1606 m
Mt. Washington 1917 m
Laurentian Plateau
Canadian Shield
Peninsule d'Ungava
Shallwood Reservoir
Newfoundland
Mt. Shasta 4322 m
Great Salt Lake
Wheeler Peak 3982 m
Great Basin
Denver
Colorado
Mt. Elbert 4402 m
Colorado Plateau
Ozark Plateau
Arkansas
Oklahoma City
Phoenix
Las Vegas
Nevada
San Francisco
San Jose
Mt. Whitney 4421 m
Los Angeles
San Diego
Coast Ranges
Sierra Nevada
Hawaiian Islands (US)
Tropic of Cancer
Guadalupe (Mexico)
Memphis
Atlanta
Mt. Mitchell 2037 m
Appalachian Mountains
Coastal Plains
Bermuda (UK)
Jacksonville
Miami
Tucson
El Paso
Rio Grande
Brazos
Dallas
Edwards Plateau
Houston
San Antonio
Gulf of Mexico
THE BAHAMAS
NASSAU
HAVANA
CUBA
Santiago
DOMINICAN REP.
HAITI
PORT-AU-PRINCE
Cayman Islands (UK)
Puerto Rico (US)
SANTO DOMINGO
ANTIGUA & BARBUDA
DOMINICA
ST. LUCIA
GRENADA
TRINIDAD & TOBAGO
PORT-OF SPAIN
JAMAICA
KINGSTON
Yucatán Channel
Yucatán
MEXICO
MEXICO CITY
Sierra Madre Occidental
Sierra Madre del Sur
Mexican Plateau
Sierra Madre Oriental
Islas Revillagigedo (Mexico)
RELMOPAN
BELIZE
GUATEMALA
GUATEMALA
HONDURAS
TEGUCIGALPA
SAN SALVADOR
EL SALVADOR
NICARAGUA
MANAGUA
SAN JOSE
COSTA RICA
PANAMA CITY
PANAMA
Golfo de Panamá
LEGEND
Country Boundary
River
Major City
Other City
Country Capital
Lake
Mountain Peak
0 500 1000 Km
0 250 500 Miles
Copyright © 2013 www.mapsofworld.com
(Updated on 18th January, 2013)
Equator
165° -180° -165° -150° -135° -120° -105° -90° -75° -60° -45° -30° -15° 0°
60° 45° 30° 15° 0°



Cordilleran orogenic belt

Extends laterally to about 6000 kilometers (From Mexico to Canada)

Approximately around 50 Ma, tectonic activity from the orogenic belt created the Laramide Rocky Mountains

Often used as an example for typical subduction related mountain building processes

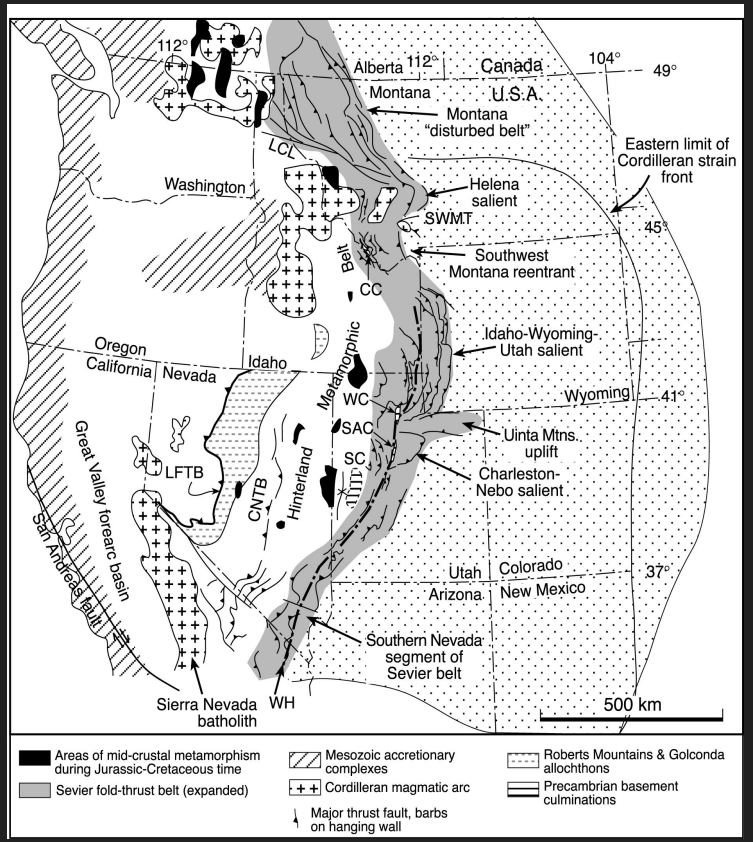
San Andreas fault

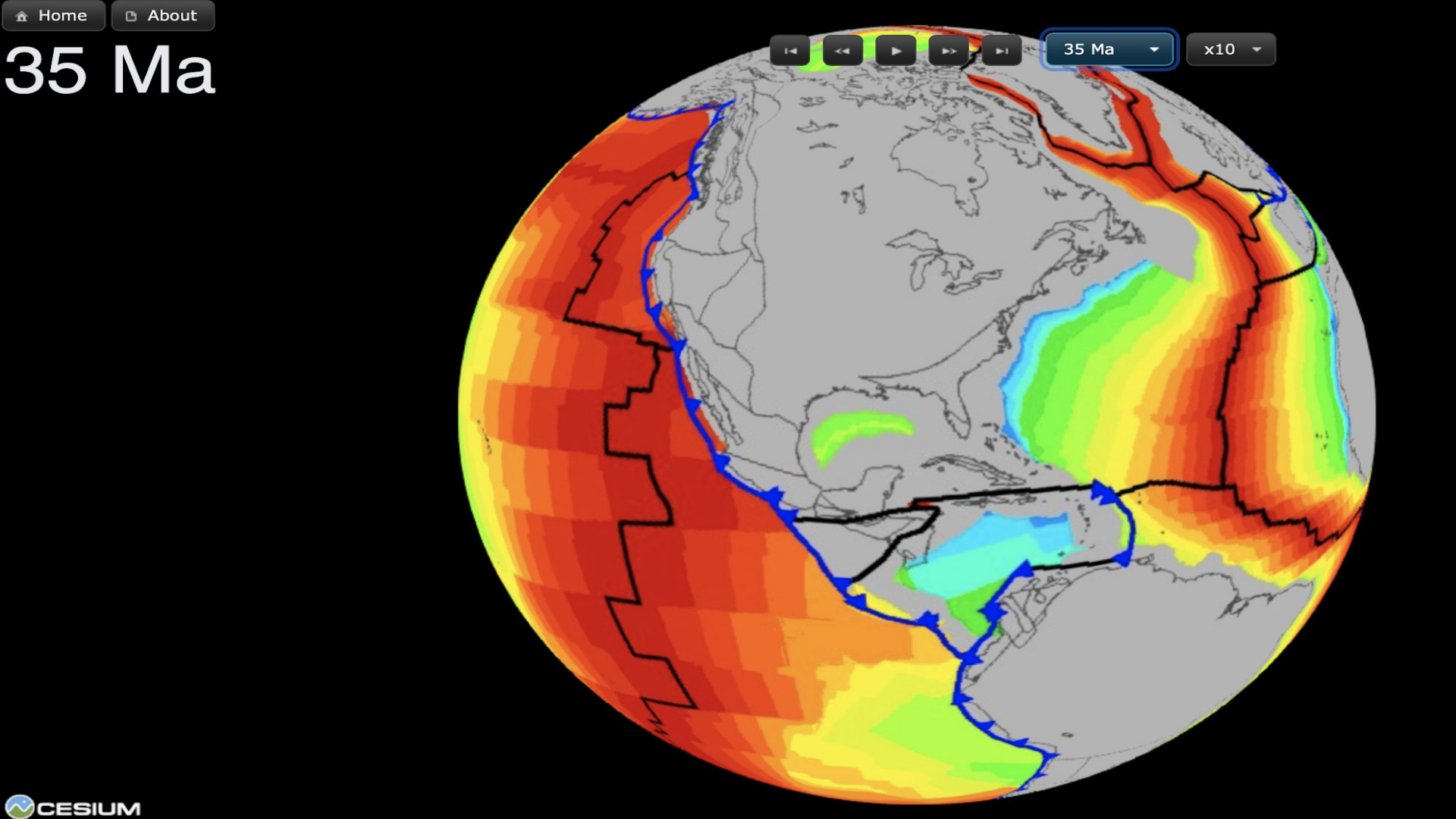
Second Tectonic event in the western coast of U.S.A.

Occurred approximately around 30 Ma.

Right - Lateral Strike Slip Fault

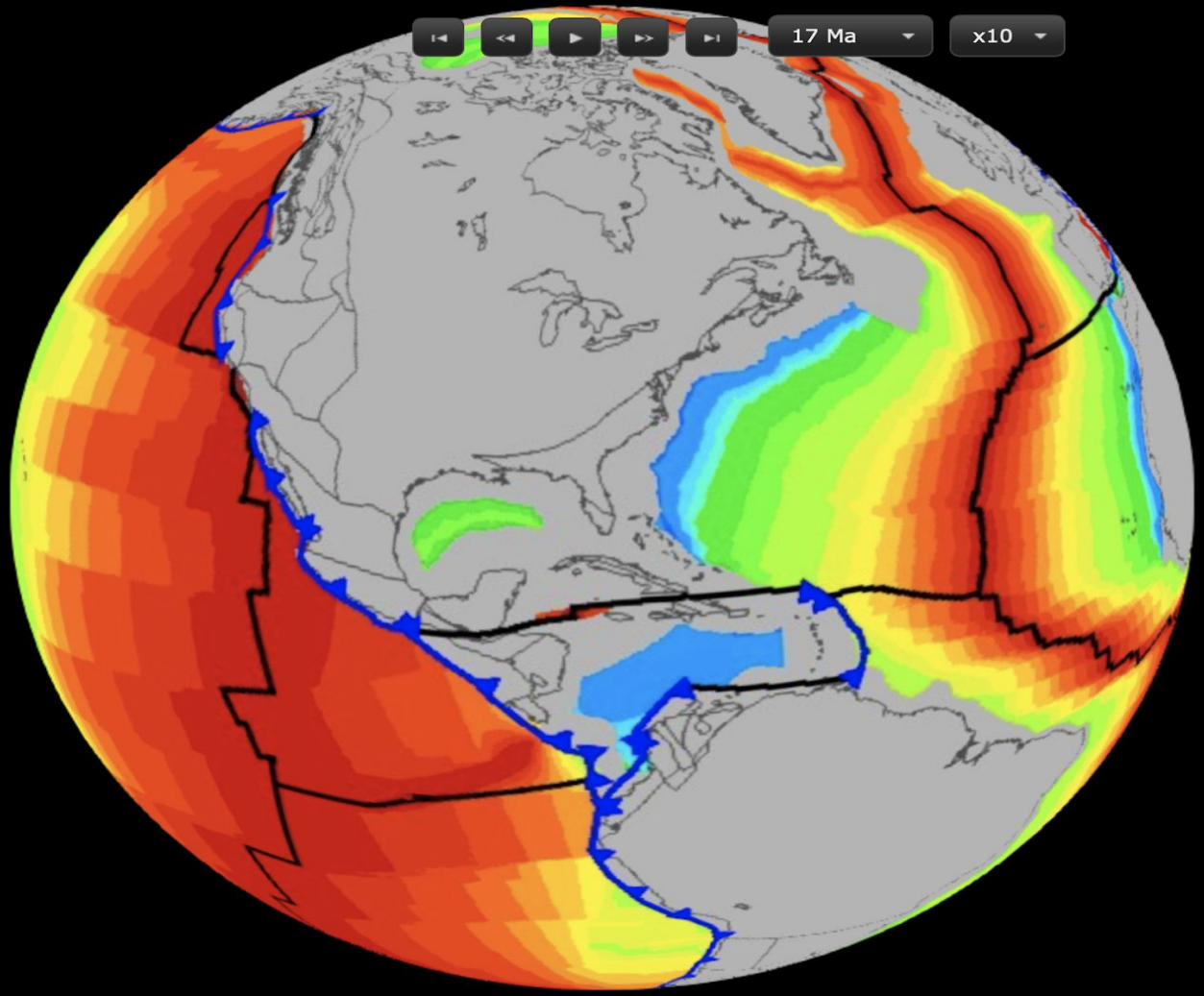
No more subduction processes occuring past this point.



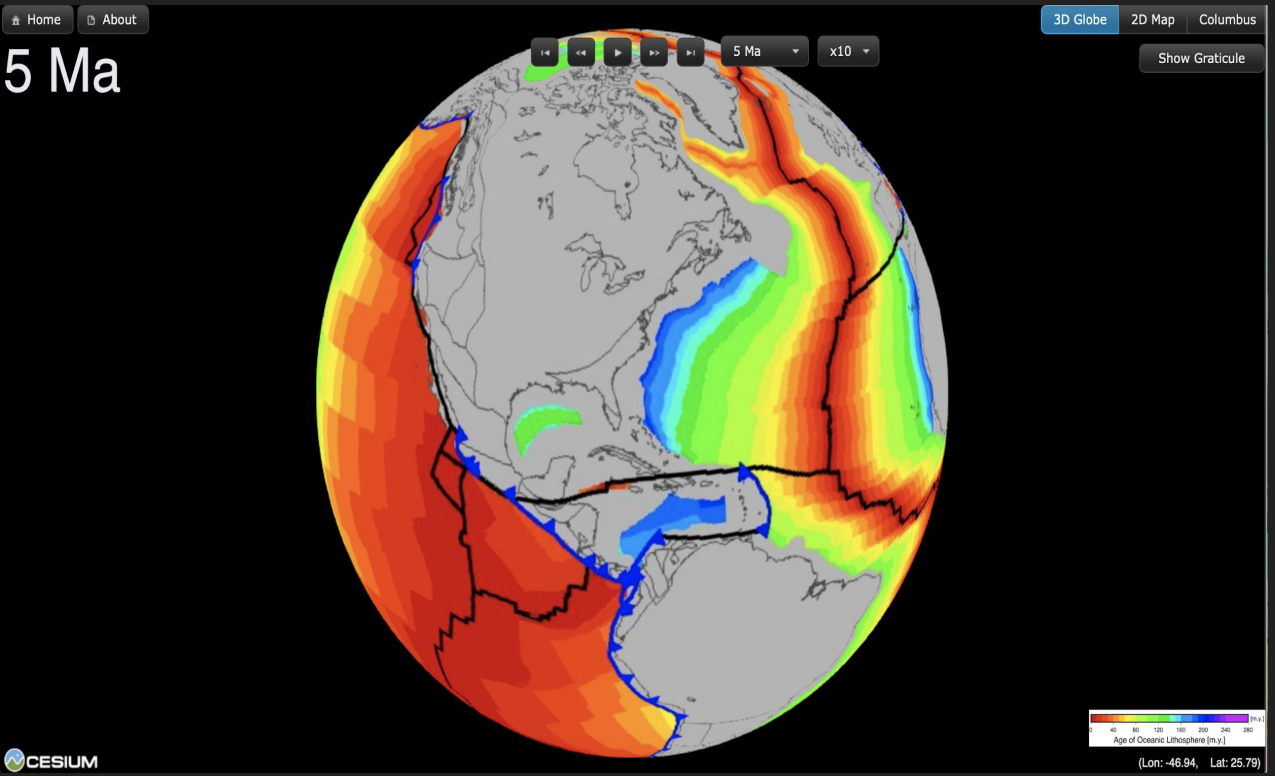

Home
About
35 Ma
35 Ma
x10
CESIUM

Home About

17 Ma



CESIUM

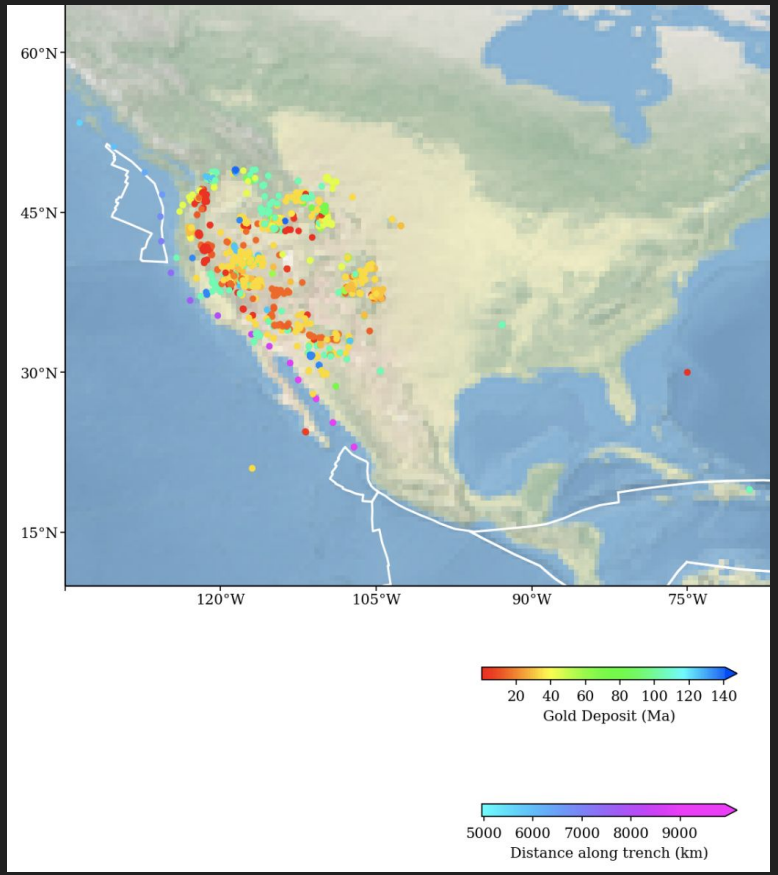


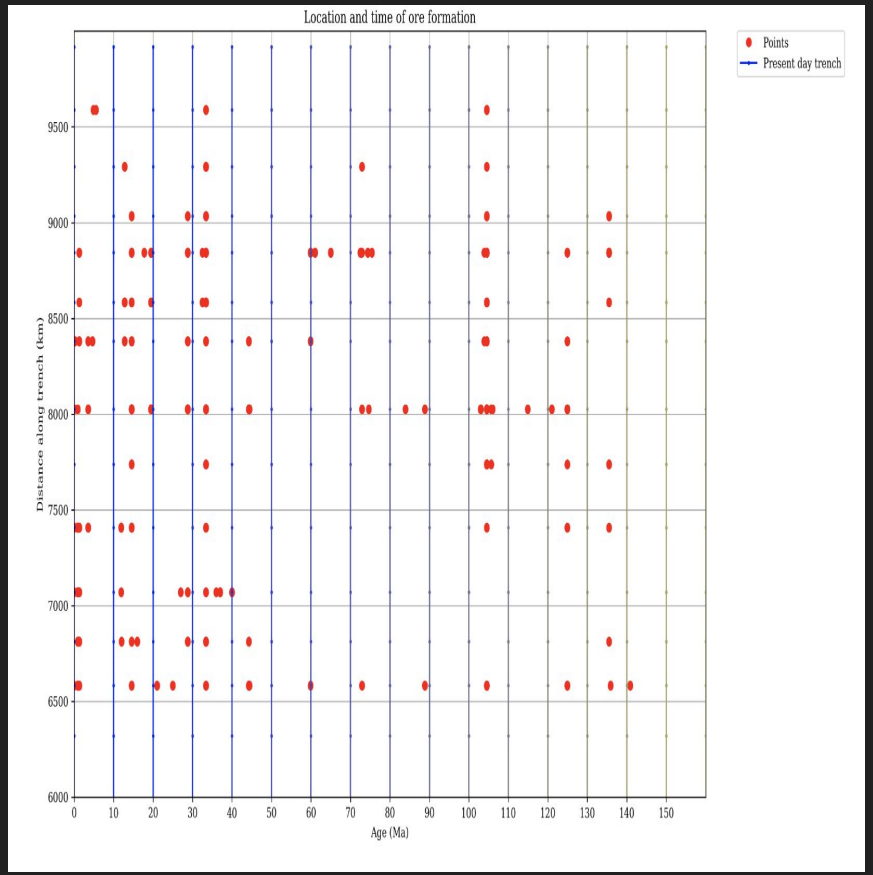
Now that we know a bit more

Let's watch the video and view the process over time

http://portal.gplates.org/cesium/?view=AgeGrid

Cross-examination of Age related deposits

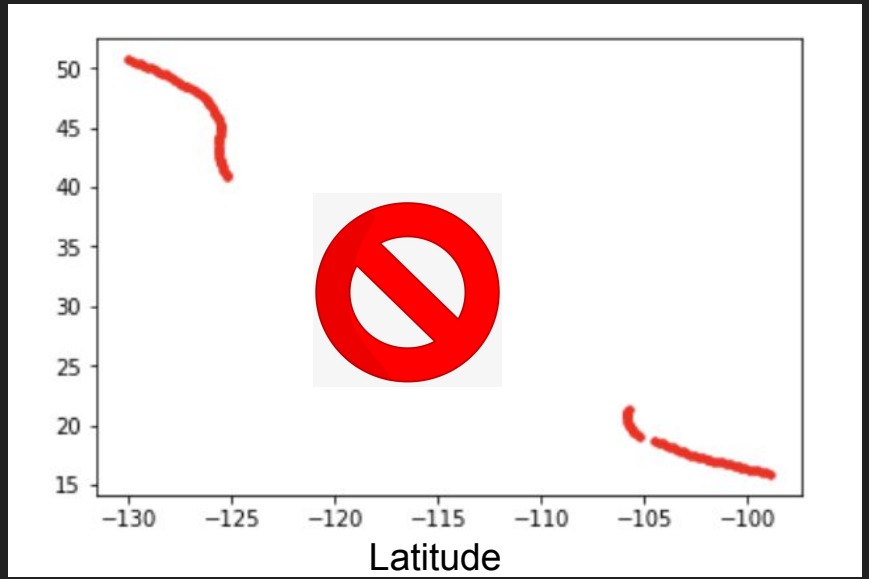




Python Method

- Pulling subduction zone point for North America from kinematic file ~ Muller_convergence/subStats_0.csv
- Plate ID = 101

(source: Earthbyte group)



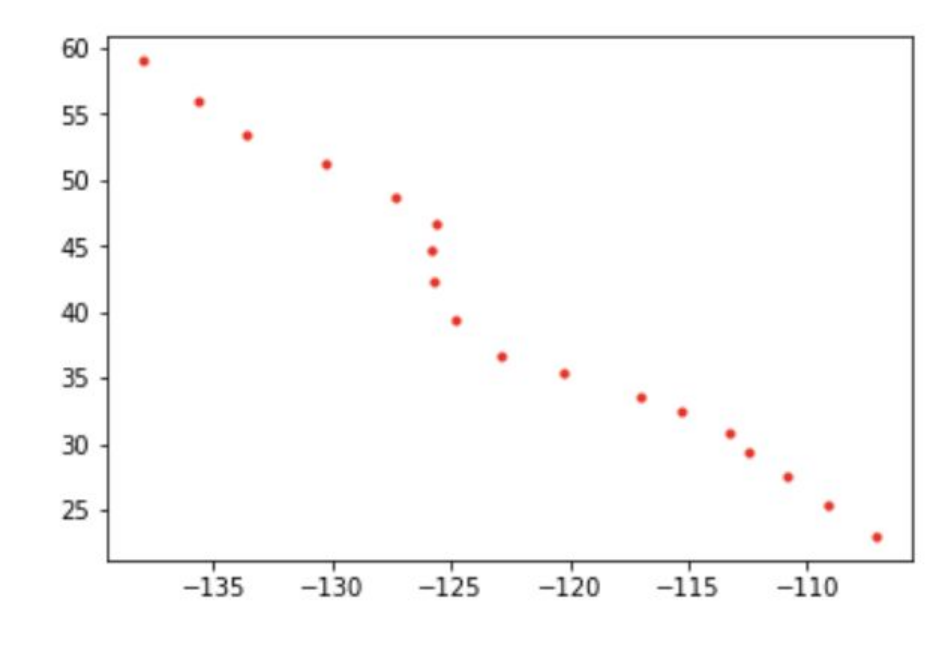
- No subduction zone!
- Causing empty data in the array

Subduction zone point

```
pointlist=numpy.array([[-137.960103,59.040862],
  [-135.672401,55.959116],
  [-133.603103,53.398991],
  [-130.284352,51.168879],
  [-127.314699,48.750958],
  [-125.639307,46.671816],
  [-125.795614,44.610062],
  [-125.712590,42.290853],
  [-124.787703,39.334025],
  [-122.933761,36.742261],
  [-120.271123,35.319573],
  [-117.027312,33.572405],
  [-115.296019,32.453982],
  [-113.289848,30.878547],
  [-112.505502,29.293555],
  [-110.789818,27.527513],
  [-109.162265,25.297241],
  [-107.134607,22.998564]])
```

- Manually input the subduction data along the current margin

- We have to do this or else it is going to cause problem when running the machine learning ‘notebook’



IT'S FIXED!

But not really...

Reading the data from txt file

- The txt file of gold deposits is in format of :

  Latitude, Longitude

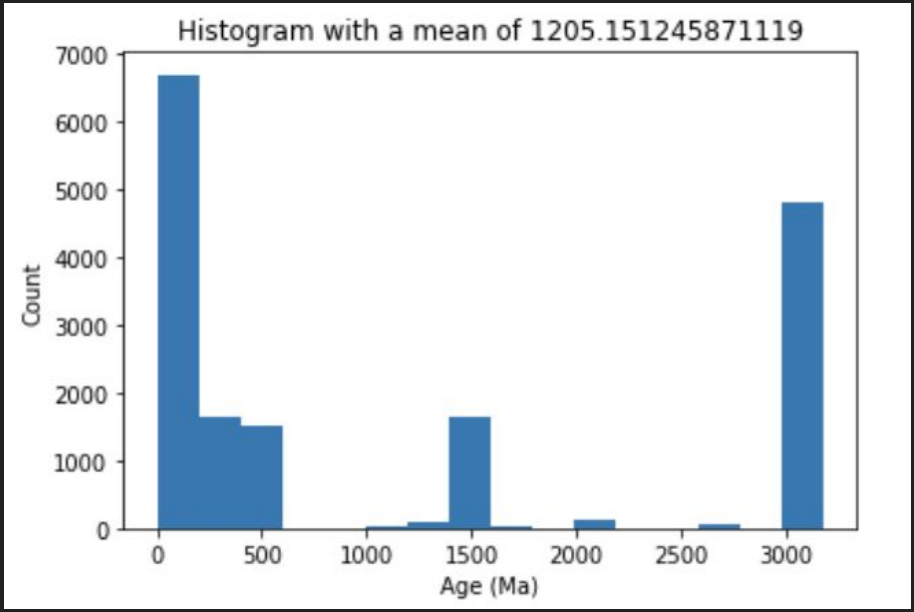
- But we are expecting :

  Longitude, Latitude

- So we have to swap their columns!

```
latlon=audata6[:,0:2]
lat = audata6[:,0]
lon = audata6[:,1]

lonlat = latlon.copy()
lonlat[:, 0] = lonlat[:, 1]
lonlat[:, 1] = lat
```



Histogram of gold deposits count in North America with respect to age

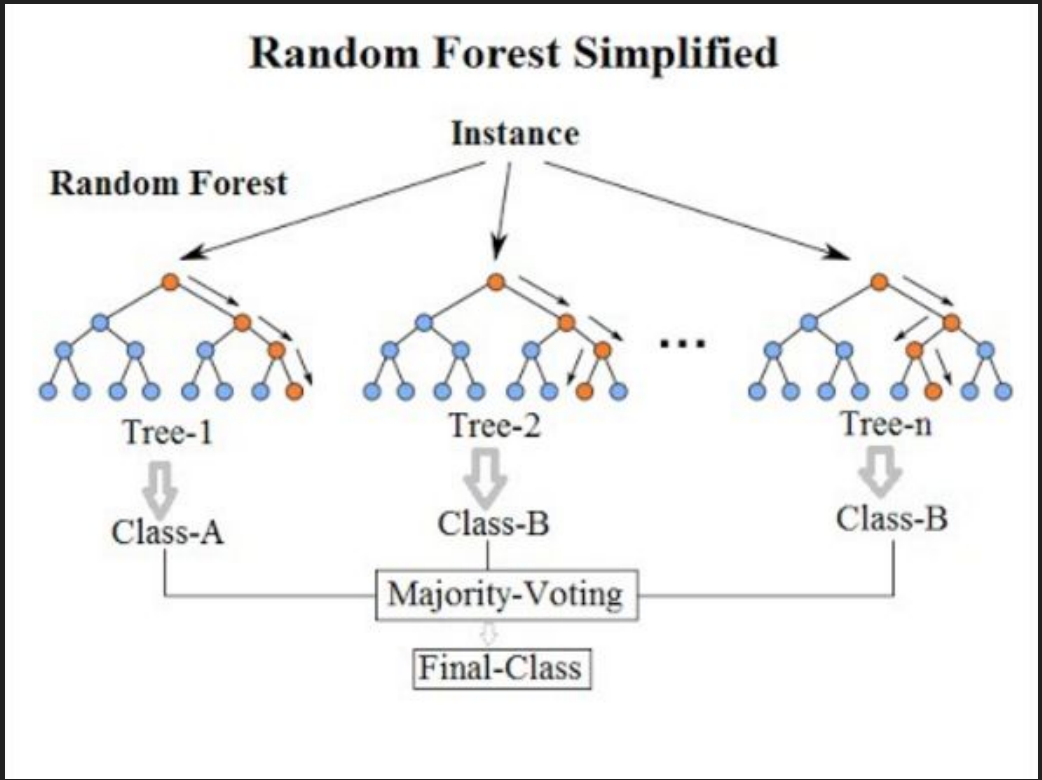
- The data is too broad (A LOT!)
  - Over 16000!
- We start by restricting the age from 0 Ma to 150 Ma
- Removed the unnecessary data for our research

Machine Learning

- Data wrangling
  - Reconstructing the geographical location of the data by ‘coregister’ the kinematic (pickle files) that associated with ore deposit at that time in that location
  - Put in a format that scikit.learn environment can read

- Data partitioning
  - Splitted into ‘training’ and ‘testing’ points
  - For training purpose by using the classifiers to later be tested with testing data for prediction
  - To determine which parameters are related to deposition and non-deposition

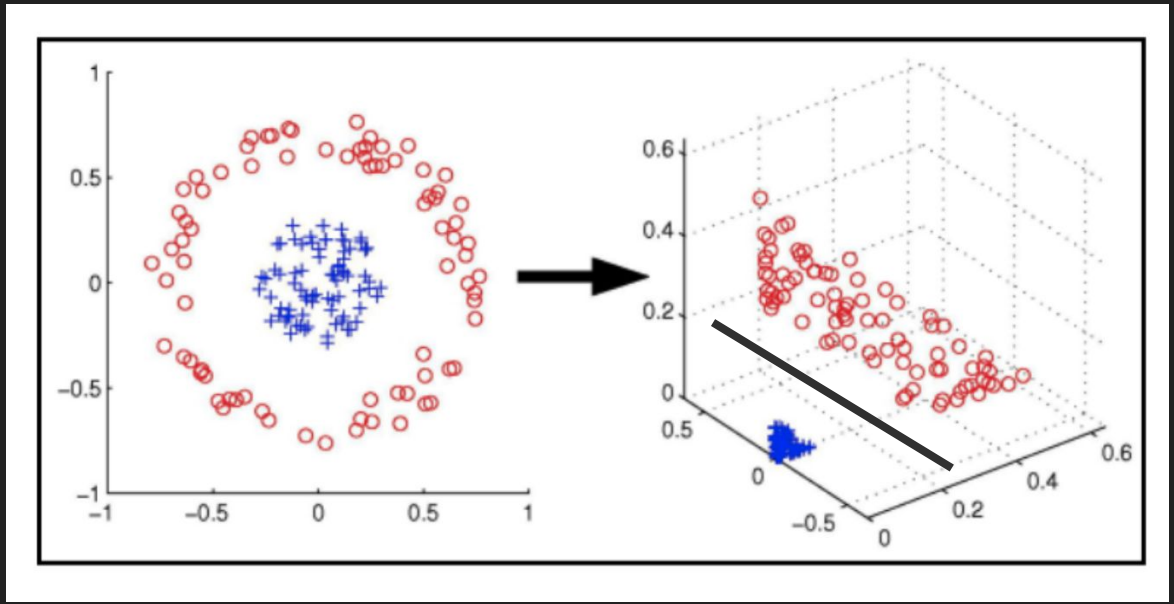
Random Forest



- Train the data set on many random decision trees

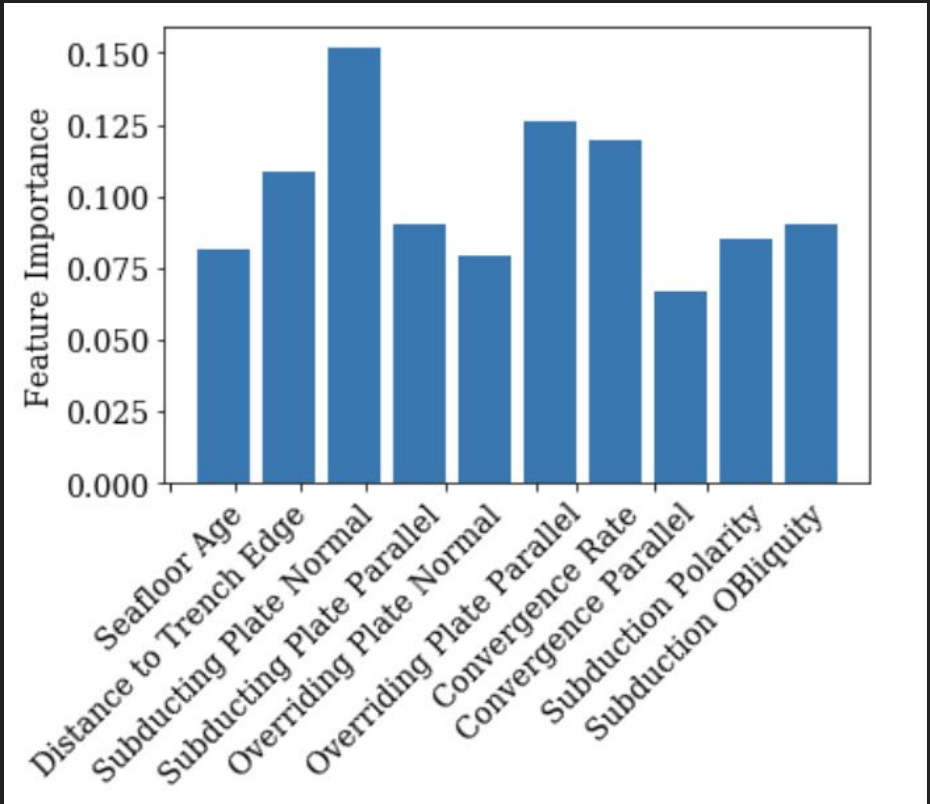
- Use “majority voting” to determine which parameters are important for prediction

Support Vector Machine



- Impossible to separate the ore deposits and non-deposits with a line
- Using ‘kernel’ trick to separate the data set
- Separate using hyperplane in 3D

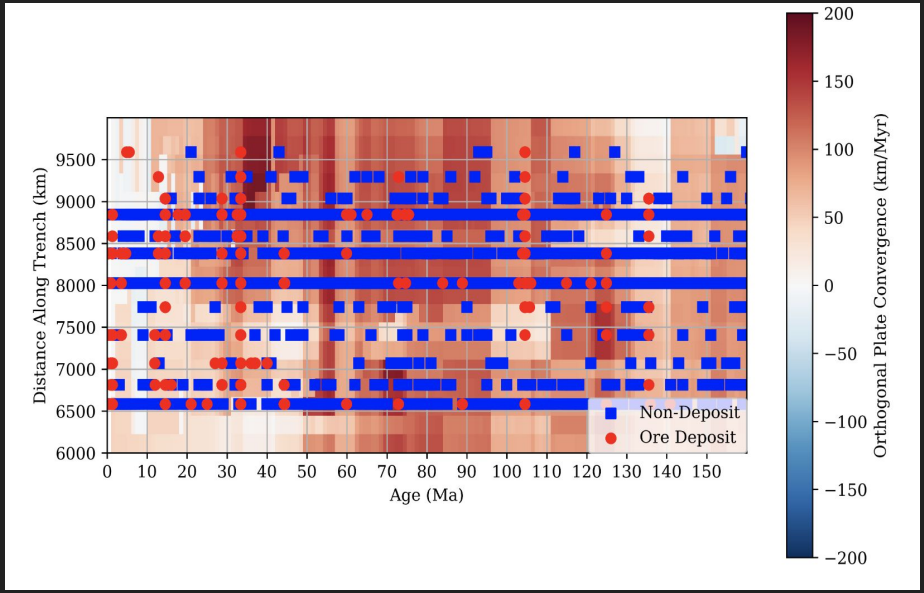
Optimised Rf Parameters



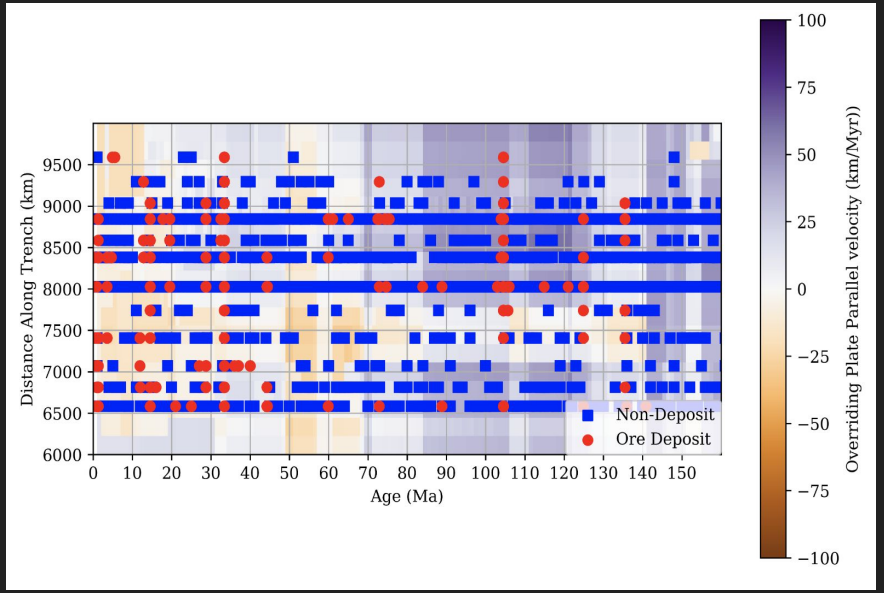
From a total of 20 parameters, Random forest was able to test and train the relationships between the parameters and the deposit data.

The 4 optimised parameters for North American tectonic setting are distance to trench age, subducting plate normal velocity, overriding plate parallel velocity and Convergence normal rate.

Relationship between parameters and gold deposits

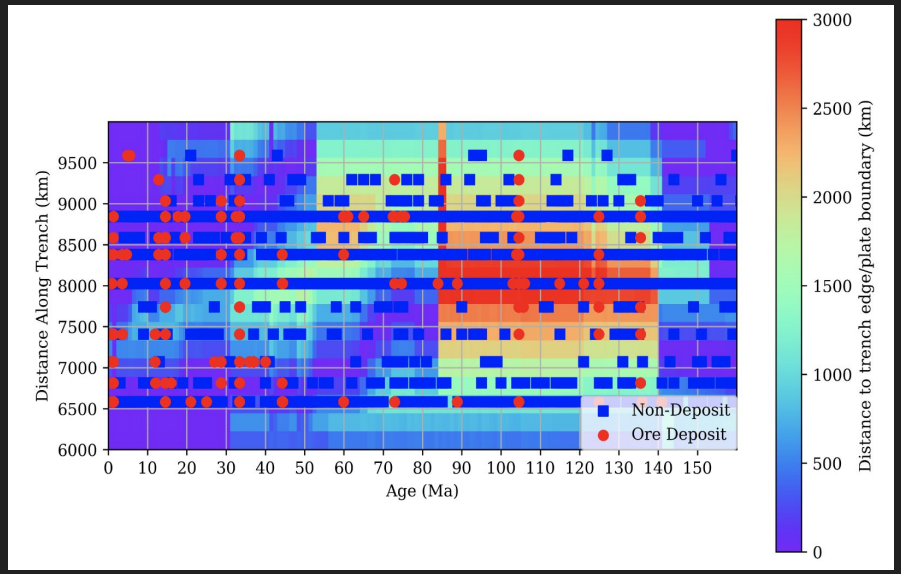


Orthogonal Plate Convergence rate (km/Myr)

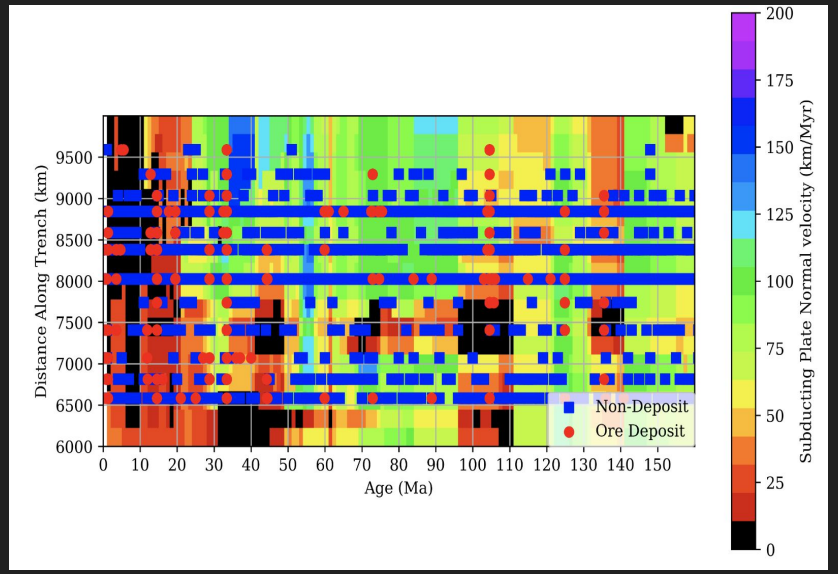


Overriding Plate Parallel velocity (km/Myr)

Relationship between parameters and gold deposits

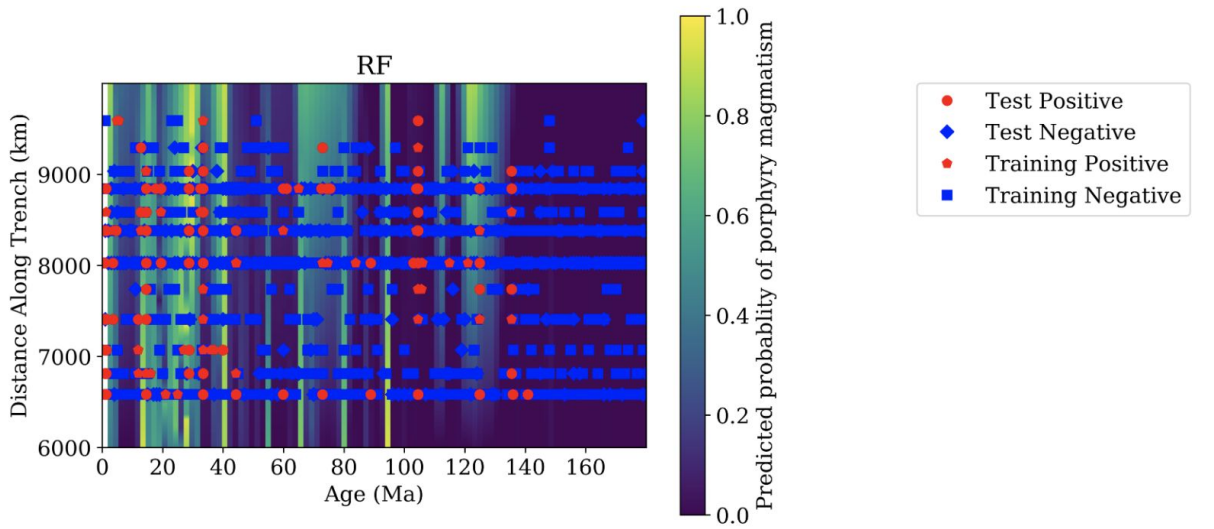


Distance to trench edge / plate boundary (km)



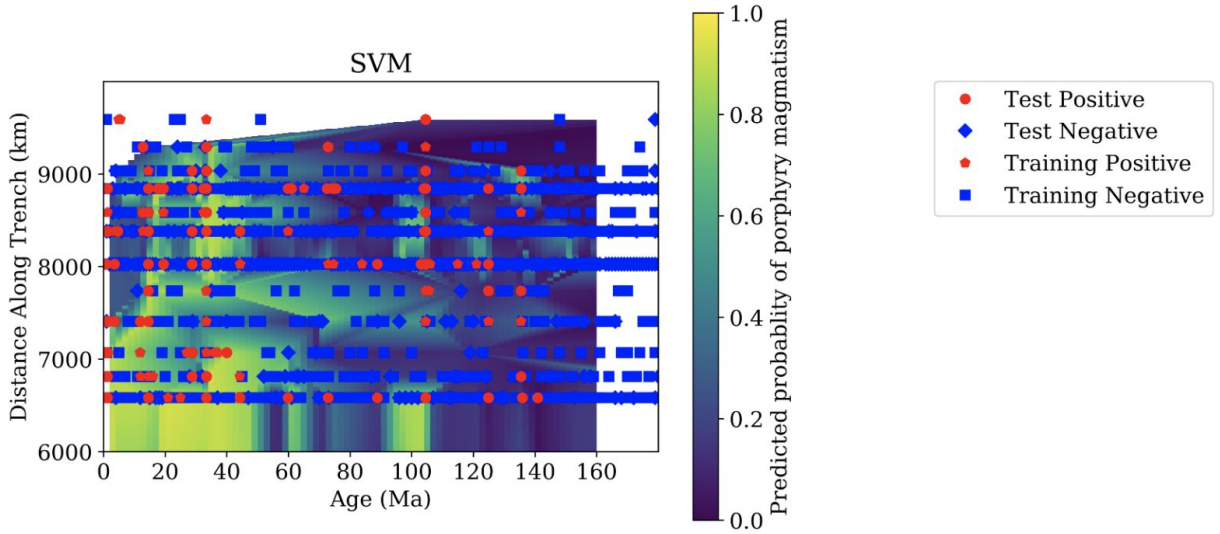
Subducting Plate Normal velocity (km/Myr)

Random Forest testing and training system



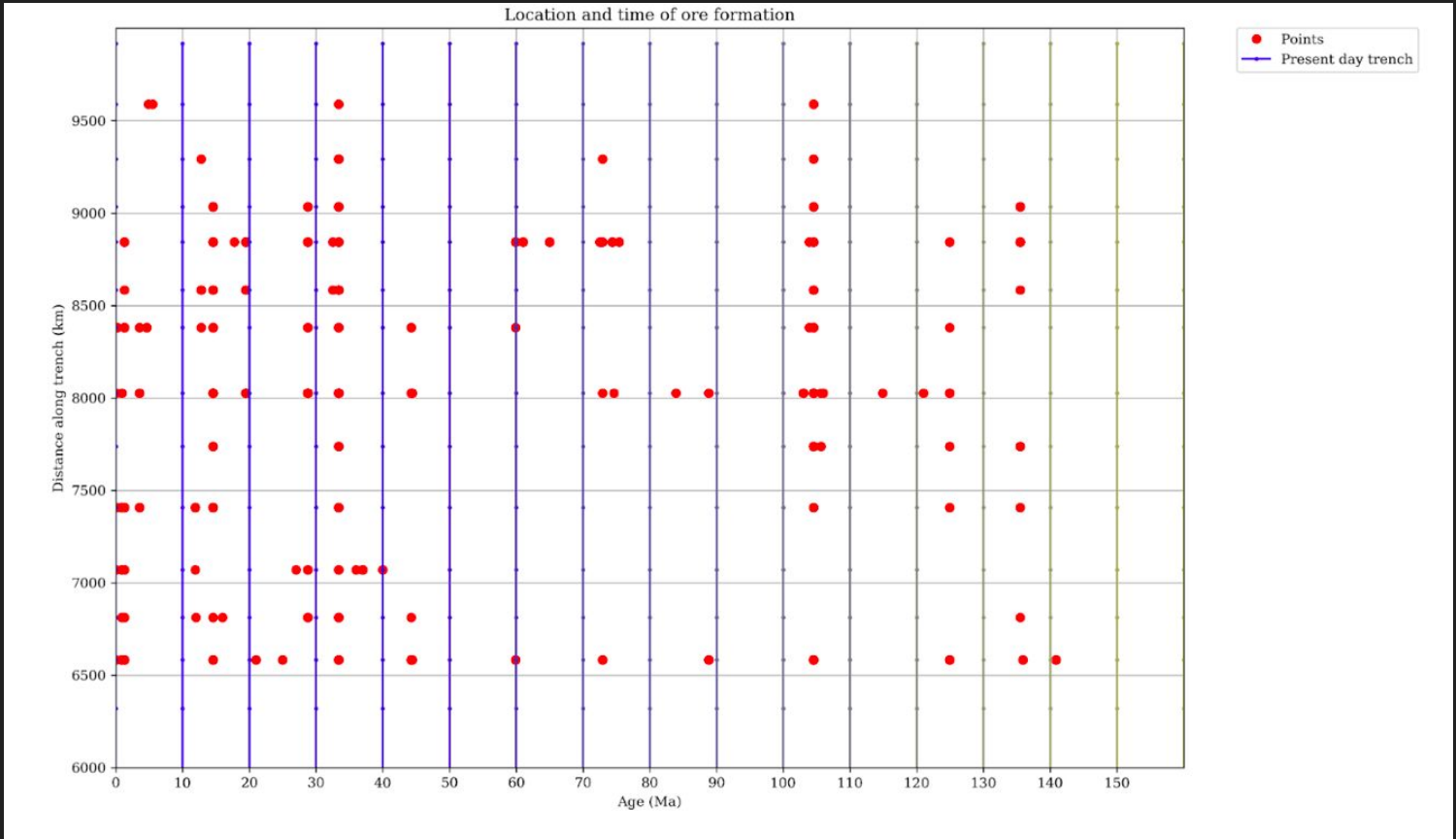

```
Five-fold cross validation scores:  [0.87146974 0.90605187 0.89740634 0.96770473 0.89965398]
SCORE Mean: 0.91 STD: 0.03
```

Support Vector Machine testing and training system




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

Limitations of dataset



Conclusion

- 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 robust nature of the provided data.
-
- The predicted results do 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 that ceased all subduction near the western coast of U.S.A

- 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.

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].

References

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

References

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).