Machine Learning Lineament Case Study: The Afar Triangle

Dan Kalmanovitch[1,2], Phil Harms[1,2], Trent Dinn[3]

[1] 4th Resource Corp. [2] GEOSEIS Inc. [3] Gno-Sys Technology Ltd.

Summary

This case study presents the results from an algorithm developed to use a machine learning classifier to detect surface structural lineaments, as expressed by topography, within the Afar triangle of Ethiopia. The Afar triangle is a region with active continental rifting and high geothermal resource potential. Surface lineaments support characterization of the deeper faults and fractures which may be associated with geothermal resources. Tectonically significant faults typically play a critical role in facilitating deep fault circulation of hot geothermal fluids.

In Harms et al. (2020) an approach to extracting surface lineament density attributes from a digital elevation model (DEM) was performed over the eastern branch of the East African Rift System, including the Afar triangle. This approach involved a sequence of geoprocessing steps and yielded adequate results at a regional scale, but not necessarily detailed enough for a geothermal site-level analysis.

A machine learning approach was investigated for detecting and mapping surface lineaments in more detail.

Theory / Method / Workflow

A machine learning model finds patterns in a set of data to make predictions with new data. To utilize machine learning for a lineament detection application, an algorithm was developed to (1) prepare geospatial data as training data, (2) implement a machine learning classifier to detect lineaments (specifically, the random forest classifier), and (3) convert the results into useful formats.

The input dataset used for building the model were extracted from the Japanese Aerospace Exploration Agency (JAXA) DEM dataset, which has a 30-meter pixel and 1 meter elevation resolution. For each input pixel, the algorithm extracted 25 elevation values from the JAXA dataset by centering a 5x5 kernel (i.e. matrix) on the input pixel. The elevations were then subtracted by the input pixel to create elevation differences. The training data for the model used the 25 elevation differences as input with a corresponding output of 1/0 (yes/no) indicating if the pixel is part of a lineament or not. The 1/0 output was extracted from a truth lineament dataset.

The truth lineament datasets were created within a test area to train the model. In this case study, a lineament is defined as a somewhat linear set of contiguous pixels following a similar azimuth. Some lineaments may be more subtle or shorter due to erosion over time and being crosscut by drainage patterns. Therefore, the truth lineaments were created to capture smaller features along with the prominent lineaments expressed within the Afar triangle region.

It was decided to prepare truth lineaments for 16 azimuth classes (e.g. N-S lineaments dipping to the east, NNE/SSW lineaments dipping to the west-northwest, etc.) instead of one truth dataset for all lineaments. By breaking the truth datasets to smaller azimuth classes, a significant improvement in the machine learning classifier predictions were observed. The accuracy of the predictions was visually assessed by comparing the predicted features with the JAXA DEM.

Results, Observations and Conclusions

Figure 1 shows an area of the Afar triangle within Ethiopia, roughly centered at 41°42’ E and 10°50’ N. Sixteen lineament azimuth predictions were created and then processed with a low-pass filter to remove some of the noise.

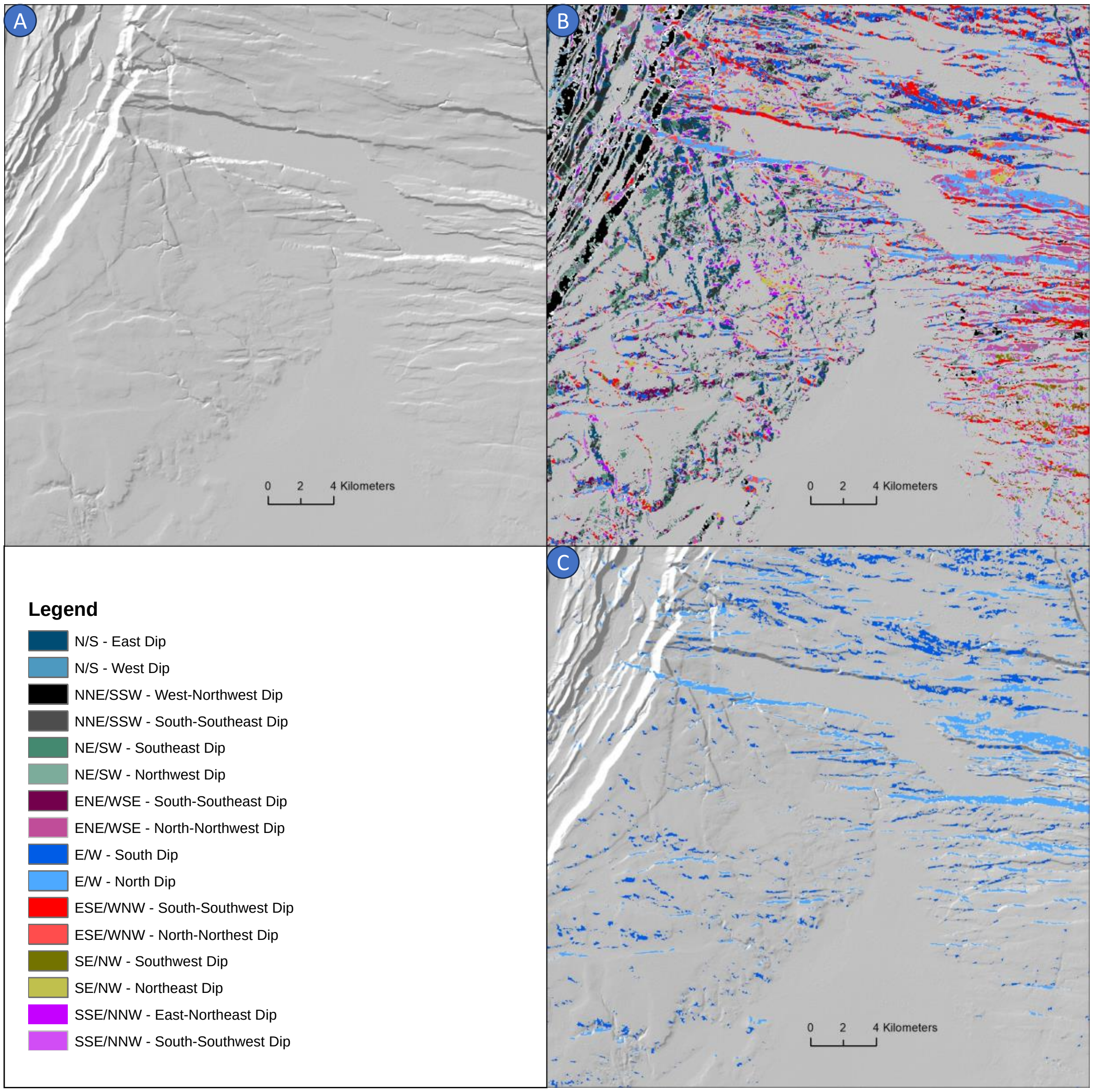

Figure 1. Lineament predictions from the machine learning classifier: (A) hillshade image derived from the JAXA digital elevation model for reference, (B) lineaments colored by their 16 azimuth classes, and (C) an example showing only E/W lineaments.

During the training process, it was learned that the quality and precision of the truth lineament datasets was very important and affected the output results significantly. In the study area, lineaments can be visually and subjectively observed on the topography. To remove input bias, a sequence of geoprocessing steps and methodologies were developed to classify the lineaments including the use of hillshade and focal statistics. During the iterations of training data, it was learned that identifying the lineament features was just as important as identifying where the features did not exist. All yes/no instances within the training area must be identified.

To improve the results, different features may be worthwhile testing in the model (e.g. using Landsat imagery), increasing the kernel size beyond 25 pixels, simplifying the number of models by implementing a multiclass classification for azimuth angle or continuing experimenting with the truth lineament datasets. Additional post-processing of the final lineaments may also suffice depending on the application. Further improvement could be possible with the implementation of deep learning models but would require a significant increase in the truth data.

Novel/ Additive Information

In this case study, due to the sensitivity of the truth lineament datasets, a fair amount of testing and experimenting was required to achieve adequate lineament predictions. By going through this process the machine learning classifier was, in a sense, testing the objectivity of the truth lineaments. Once achieved, however, the predictions have an unbiased quality to allow a geoscientist confidence in their interpretation of the data in support of geothermal resource characterization.

Training datasets and detailed work can be done over a relatively smaller area and then the machine learning process can be run efficiently over regionally extensive areas generating high quality, useable surface lineament datasets based on topography.

References

Harms, P., Kalmanovitch, D., Birnie, D., Hickson, C.J. "An Innovative Approach to Geothermal Prospecting through Interactive Analytics". Proceedings, 8th African Rift Geothermal Conference, Nairobi, Kenya: 2 – 8 November 2020