University of Zurich[UZH]

Master Thesis[1]

Mining Related Conflicts and Cooperations in Peru – To build an event dataset

Author:
Zhang Jingwen (17-742-487)
Master's degree programme with major in Political Science (90 ECTs)
Department of Political Science
Minor in Geographic Information Science (30 ECTs)
Department of Geography
University of Zurich

Supervisor:
Prof. Dr. Thomas Bernauer
Institute of Science, Technology and Policy
ETH Zurich - Department of Humanities, Social and Political Sciences

Co-Supervisor:
Megan Seipp, doctoral candidate
Institute of Science, Technology and Policy
ETH Zurich - Department of Humanities, Social and Political Sciences

Zurich, 31st of May 2021

[1] This version is the spatial analysis part of my master thesis. Since I have signed a non-disclosure agreement with the institute, please do NOT circulate publicly without notifying the researchers.

Abstract

The mining industry is essential for the global economy. In the past 15 years, the number of reported conflicts related to the mining industry increased rapidly. In Peru, mining conflicts are mainly associated with water scarcity, pollution, past conflicts and distribution of mining revenue. To mitigate the increasing conflicts and avoid losses, more and more companies start to incorporate corporate social responsibility [CSR] into their operation. In this thesis, I built an event dataset to monitor conflicts and cooperation and their geographical distribution from 2009 to 2019 to find out the main drivers of conflicts and the effects of cooperation. The event dataset contains 432 mining-related events from 2009 to 2019. According to the results of the analysis of the dataset, I find that copper mines have the most conflicts and cooperations, followed by gold, silver, and zinc mines. Open-pit mines have more than double of conflictive events than underground mines. In the spatial analysis, regions with high conflict are also regions with high cooperation levels, such as regions Arequipa and Cajamarca. This is similar to findings from the event dataset that cooperative actions in Peru are reactions to existing conflicts or past grievances instead of proactive actions. By comparing the spatial distribution of events and other spatial data layers, I find that the denser the population is, the more conflictive a mine site was reported by the news; the more impoverished the local population around a mine site is, the less conflictive the mine site is. In addition, if a mine site is located at a higher altitude, the less conflictive it is.

1.1 Spatial analysis – data and method

In this part, data and methods to visualize the geospatial distribution of information captured by the event dataset and to test hypotheses (H1, H2, H5, H6) are given.

1.1.1 *Geospatial distribution of MES*[2]

Data: The map of Peru is downloaded from the Database of Global Administrative Areas [GADM]. It includes four different levels of administrative areas. The first level is the polygon of the whole country, which will be used as the mask to extract other global geospatial layers for Peru. The second level consists of 26 polygons, including 25 regions and the Lima Province. The regions are further divided into 196 provinces. So there are 196 polygons at the third level. Regions in Peru are similar to states in the USA and cantons in Switzerland. At the fourth level, provinces are consist of 1869 districts.

Method: To show the overall distribution of events in Peru, the second level of administrative areas (25 regions and Lima Province) is selected. MES values for all events are summed up in each region to capture both MES intensity and event frequencies. To plot the summed MES values (also called MES scores in the later part), I joined this table to the attribute table of the region polygons based on the same names of regions. For visualization, MES scores are classified into seven classes by the method "natural break (Jenks)" and colored by a diverging color gradient.

1.1.2 *Population density and poverty rate*

As discussed in the previous part "theoretical arguments and hypotheses", two hypotheses are derived from the social movement theory about resources for mobilization. In this part, I will use spatial data to test these two hypotheses:

[2] MES [Mining Event Scale] indicates the intensity of mining-related events on an 11-level ordinal scale, ranging from -5 (most conflictive event) to +5 (most cooperative event). Value 0 is assigned to neutral events. Those events are collected from media reports from 2009 to 2019 by the author, and they are coded with geo-coordinates. All information is joined in the event dataset.

H1: The denser the population where a mine site is located, the more conflictive the mine site is.

H2: The more impoverished the local population where a mine site is located, the less conflictive the mine site is.

1) Population density

Data: The UN WPP-adjusted population density rasters are retrieved from the Center for International Earth Science Information Network [CIESIN]. It is a data center in NASA's Earth Observing System Data and Information System [EOSDIS]. The data is available for the years 2000, 2005, 2010, 2015 and 2020. The year 2015 is chosen because this is around the middle of the research timespan here. The raster dataset consists of estimates of the number of persons per 1 km^2, based on counts consistent with national censuses and population registers for relative spatial distribution, but adjusted to match the 2015 revision of the United Nation's World Population Prospects country totals. The rasters were adjusted by dividing the UN WPP-adjusted population count raster for a given target year by the land area raster. The rasters were produced at 30 arc-second horizontal resolution, approximately 1 km at the equator (CIESIN 2018).

Method: The global dataset is downloaded in GeoTiff format and extracted by masking the polygon of Peru to retrieve the population density map for Peru. The spatial analysis in this part is mine project-based. There are in total 77 mine projects and illegal mining activities in the region Madre de Dios, namely the event cluster "illegal gold production". Among these 77 mine projects, some of them only have one event. I excluded mine projects with only one neutral event because those cannot show problems at the current stage. After excluding those, in total, there are 67 mine projects and one cluster of illegal gold production. The MES values of them are summed up as MES scores. After plotting these 68 points data to the population density map, the population density of each point is retrieved by the function "Extract Values to Points" processed by ArcGIS Pro. Since there are only 68 observations, it is not suitable to run a regression analysis to test the relationship between population density and MES scores of each point. Instead, a scatter plot and the linear trend line between the two variables are fitted and processed by ArcGIS Pro functions.

2) Provincial poverty rate

Data: The tabulated poverty rates are derived from the National Institute of Statistics and Informatics [INEI] of Peru. The poverty rate of 2009 is available at the level of regions, provinces, and districts. Since there are too many districts (in total 1869) in Peru, the level of provinces is selected to produce the poverty map. INEI produced the data by combining household surveys and the census to evaluate the poverty rate at each district (Ministry of Economy and Finance of Peru 2021a). The poverty rate indicated the percentage of the population in the district/province/region living in poverty. According to the measurement of the monetary poverty rate by INEI, a household is said to be poor when its per capita spending is below the Total Poverty Line (LPt). The total poverty line varies across regions since the commodity prices are different in different regions, but it ensures 2318 Kilocalories per day per capita. If the consumption cability of households is below it, it is considered poor (Ministry of Economy and Finance of Peru 2021b).

Method: The third level of administrative areas (1869 provinces) is selected. I joined the table of MES scores for 68 mine sites to the attribute table of the province polygons based on the same names of provinces. To simplify the visualization, the provincial poverty rate is classified into 7 classes by the method "natural break (Jenks)". Since the provincial poverty rate is a polygon layer, instead of using the function "Extract Value to Points" as before, the function "spatial join" is applied to join the point feature of mine sites' MES scores and the polygon feature of the provincial poverty rate.

1.1.3 Elevation and water resources

According to the theoretical argument of socio-environmental characteristics and livelihoods around mine sites, two hypotheses are derived from the previous "theoretical arguments and hypotheses". These are:

H5: The higher altitude a mine site is located, the less conflictive the mine site is.

H6: The scarcer the water resource where a mine site is located, the more conflictive the mine site is.

1) Elevation

Data: The elevation raster is retrieved from the Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer [ASTER] Global Digital Elevation Model [GDEM] Version 3 [ASTGTM], via NASA's EARTHDATA portal. The DEM has a spatial resolution of 1 arc-second (approximately 30-meter horizontal posting at the equator). The ASTER GDEM (Version 3) data were created from the automated processing of the entire ASTER version Level 1A archive of scenes acquired from 2000 to 2013 (LP DAAC 2021).

Method: Since the original file is enormous due to the high spatial resolution, it is downloaded by dividing Peru into two geographical segments. The raster data is downloaded in GeoTIFF format and extracted by masking the polygon of Peru to retrieve the DEM for Peru. In ArcGIS Pro, both of them were merged by the function "Mosaic to new raster". Since the merged raster is enormous and takes a long time to process, a new raster with a coarser resolution (500 m) is created by the function "resample" with the bilinear interpolation method. The original merged DEM is for extracting elevation values for these 68 mine sites. The coarser DEM is for producing the output map. The elevation values are visualized by the method "stretch (standard deviation)" and colored by a default gradient for elevation. Other analysis methods are conducted as the same in the population density map.

2) Water resources

Data: In the research by Salem et al. (2018), they used rainfall as a proxy for water competition and access to water. They found out that lower rainfall is associated with more water-related conflicts. In this thesis, I used the rasters of precipitation (Version

1.2) from the climate datasets "Climatologies at high resolution for the Earth land surface areas" [CHELSA]. It is about the average annual precipitation (mm/year) from 1979 to 2013. The rasters were produced at 30 arc-second horizontal resolution, approximately 1 km at the equator (CHELSA 2021).

Method: The global rasters dataset is downloaded in GeoTiff format and extracted by masking the polygon of Peru to retrieve the annual precipitation map for Peru. The annual precipitation values of 68 mine sites are retrieved by the function "Extract Values to Points" processed by ArcGIS Pro. A scatter plot and the linear trend line between the two variables are fitted within ArcGIS Pro. The annual precipitation values in Peru are visualized by the method "stretch (standard deviation)" and colored by a default colour gradient for precipitation. Other analysis processes are conducted as the same in the population density map and elevation map.

1.2 Findings of spatial analysis

In this part, findings of the distribution of information captured by the event dataset and results of the testing of 4 hypotheses (H1, H2, H5, H6) are provided.

1) Geospatial distribution of MES

In Figure 27, the scores are the summed MES values. Except for the grey regions where no relevant event was reported, the redder the region's color, the more conflictive it is, or conflicts in the region are greater than cooperations in terms of both numbers and intensity. This also means that conflicts are far from being solved. If the color is bluer, it means conflicts and cooperations are equally canceling each other, or conflicts were mostly resolved. In general, most regions in Peru are found to have more mine-related conflicts than cooperations. The length of bars in each region indicates the number of neutral events, which happened most frequently in the capital Lima because of new policies and decisions regarding national strikes.

To compare the spatial distribution of conflicts and cooperation, MES values are summed separately. Figure 28 and Figure 29 show that regions with high MES conflict scores are also regions with high MES cooperation scores. For example, regions Arequipa and Cajamarca. This is similar to findings from the event dataset that cooperative actions are reactions to existing conflicts or past grievances instead of proactive actions such as improving the current situation and taking preventative measures. Therefore, regions with high MES cooperation scores also indicate that there are problems.

Region Arequipa has the highest MES scores for both conflict and cooperation. It has main mine projects such as Tia Maria (copper), Orcopampa (gold), Cerro Verde (copper); region Cajamarca has mine projects Yanacocha (gold) and Conga (gold). All of them are on the list of mine with most news reports, as shown in Table 5.

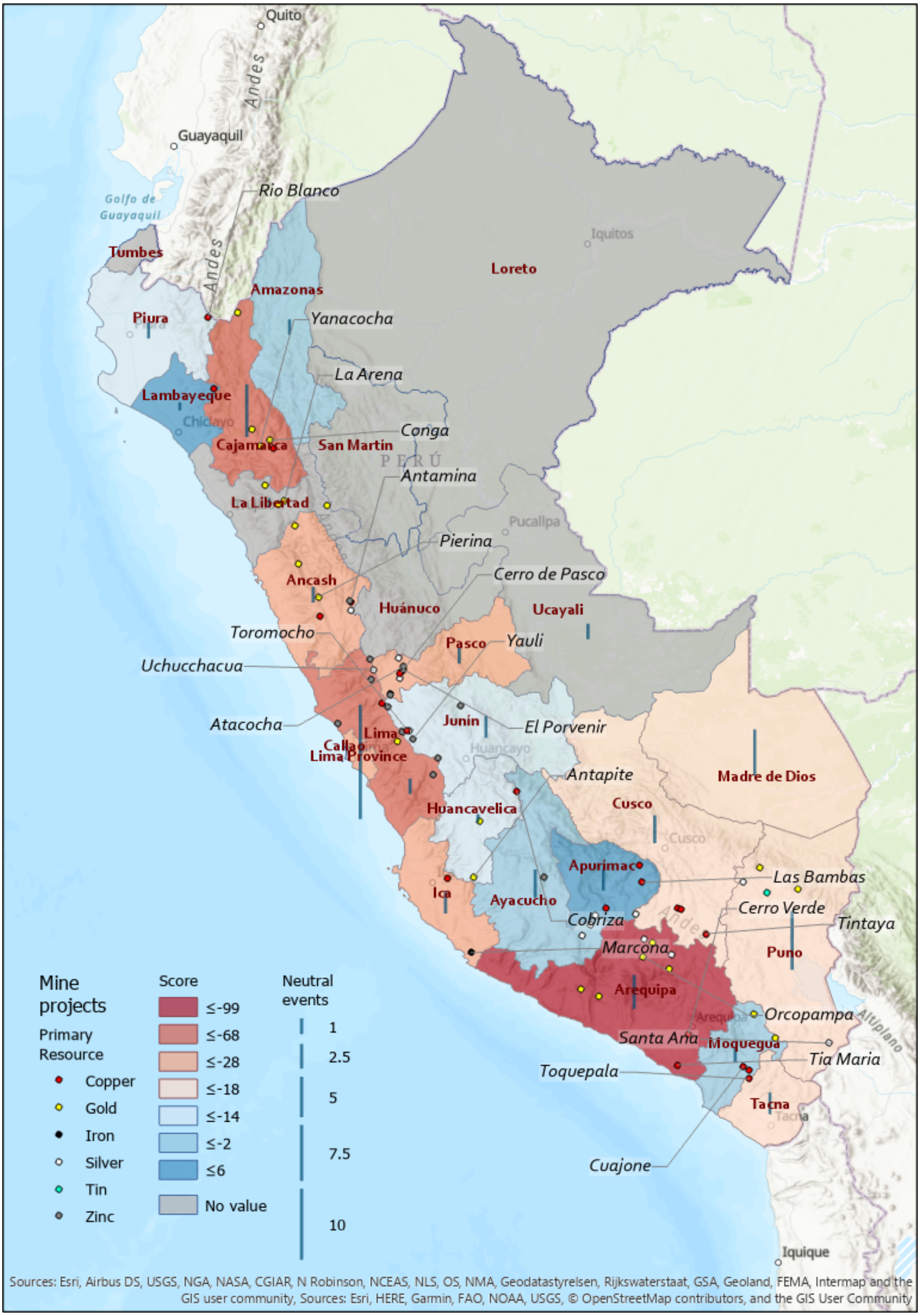

Figure 1: MES Map of Peru

Figure 2: MES Map of Peru – Conflictive events

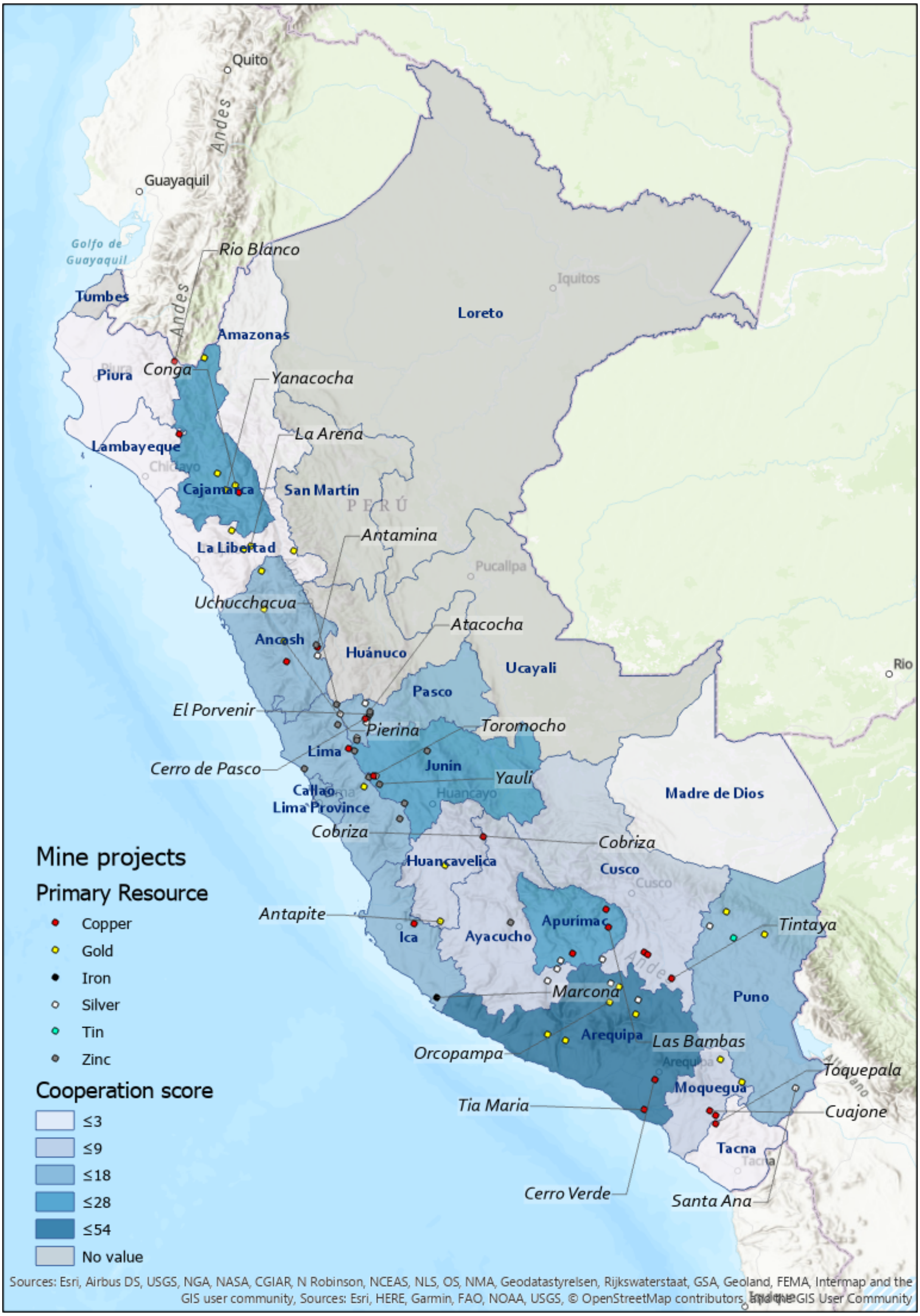

Figure 3: MES Map of Peru – Cooperative events

2) Population density

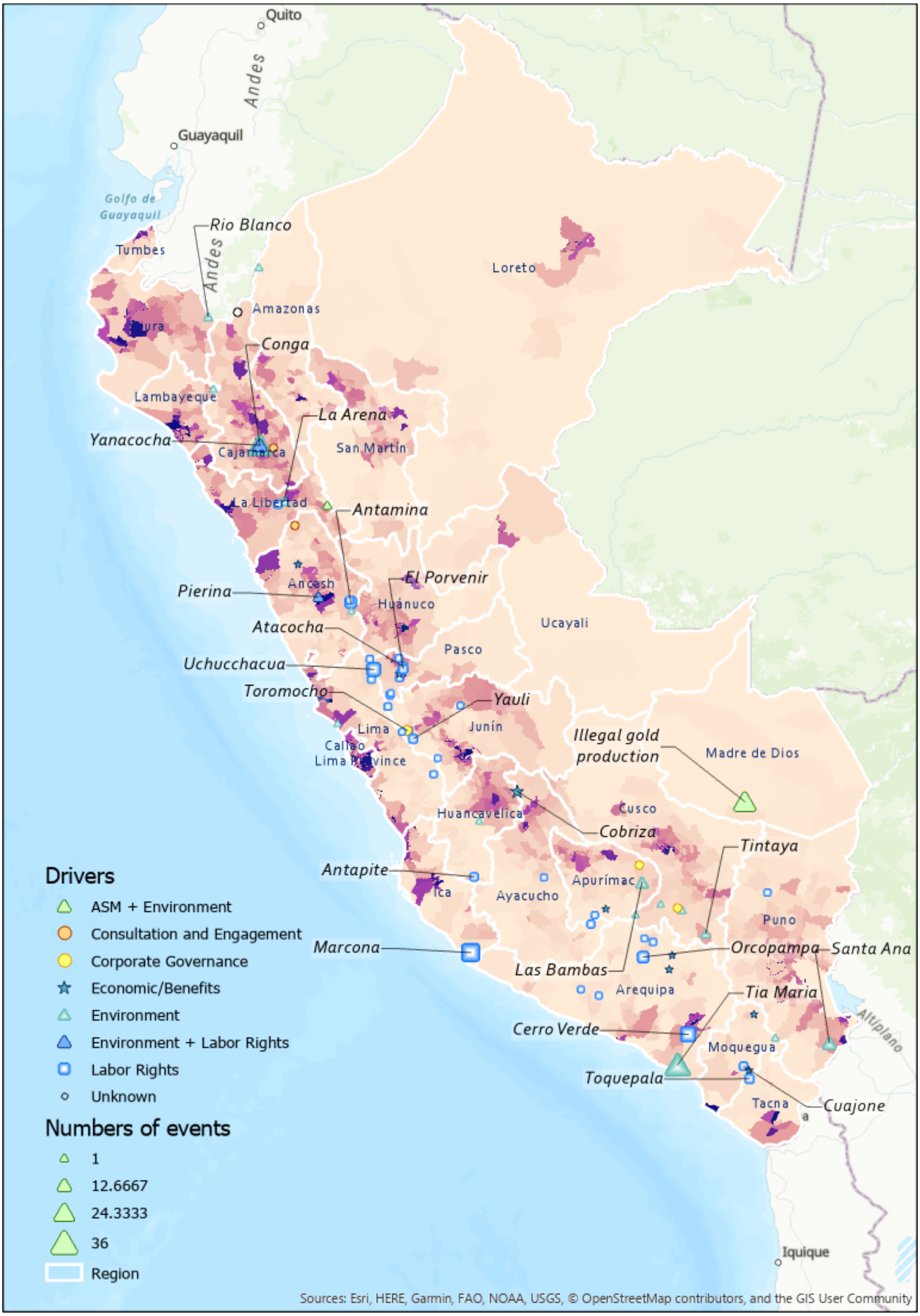

Figure 4: Population density map of Peru (2015)

As shown by Figure 30, the darker the color of the raster cell is, the denser the population is. The method used to symbolize the density is “discrete” based on 256 colors on the same color gradient. Therefore, the legend is not available because it is not possible to show all 256 levels. I didn’t choose the method such as “classify” because most regions have a sparse population, so the difference in population density varying across cells is not apparent enough for an A4-size map. The population density of Peru is ranging from 0 to 30103 persons per 1 km². The white outline is the border of each administrative region. There are roughly seven drivers behind events, and each is marked in a different shape and color. The size of the symbol indicates the number of events reported for each mine site. The bigger the symbol is, the more news were reported for the mine sites. Mine sites with more than five reported events are labeled on the map.

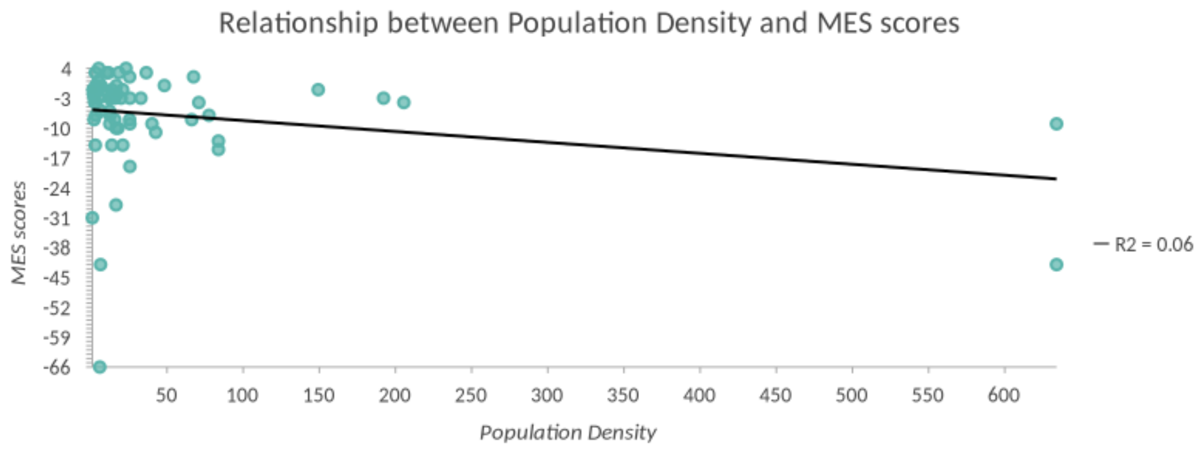

Figure 5: Scatterplot of the relationship between population density and MES scores

The scatter plot shows the relationship between population density and MES scores regarding these 68 mine sites. MES scores are the summed MES values for each mine site. The lower the MES scores, the more conflictive events were reported for the mine sites than the reported cooperative events. Although most mine sites are located in relatively low population densities, the trend line[3] shows that the denser the population is, the more conflictive a mine site was reported. It confirms the hypothesis (H1) that human as one of the essential resources of social mobilization, the higher value of it can increase social conflicts. Noticeable is that population density values are exactly

[3] Slope and intercept of the trend line: y = -5.68698 + -0.02559 x

the raster cell value where each point is located. If not considering the heavy workload, it is better to consider the average population density within 25 km or 50 km, depending on different research goals.

3) Poverty rate

The provincial poverty map of 2009 (Figure 33) indicates that most provinces in Peru were generally poor by 2009. The average poverty rate of provinces is around 48.8% in 2009. It means that around 48.8% of the whole population can not ensure 2318 Kilocalories per day per capita. The most impoverished regions are Huancavelica, Apurímac, Huánuco, Ayacucho and Puno regions, all of which have an average poverty rate higher than 60%. Since Peru has achieved a lot to combat poverty since 2009, the poverty map of 2009 is more about a relative comparison between provinces. As mentioned in the dataset analysis, together with the region Ayacucho, Apurímac is the birthplace of the Shining Path - the Communist Party of Peru. Residents in Apurímac left the region to escape poverty and political violence during Peru's internal conflict with the Shining Path insurgents in the 1980s and 1990s (Dube 2015). Apurímac used to be a region relying on small-scale agriculture. After discovering important mining resources e.g. copper, big mine projects such as Las Bambas started to operate in the Apurímac region. Managing such large mining revenues will be a new challenge to the Apurímac region.

As shown by Figure 32, the trend line indicates a positive relationship between poverty rate and MES scores. This confirms the hypothesis (H2) that the more impoverished the local population where a mine site is located, the less conflictive the mine site is, because monetary resources are also the main resources of social mobilization.

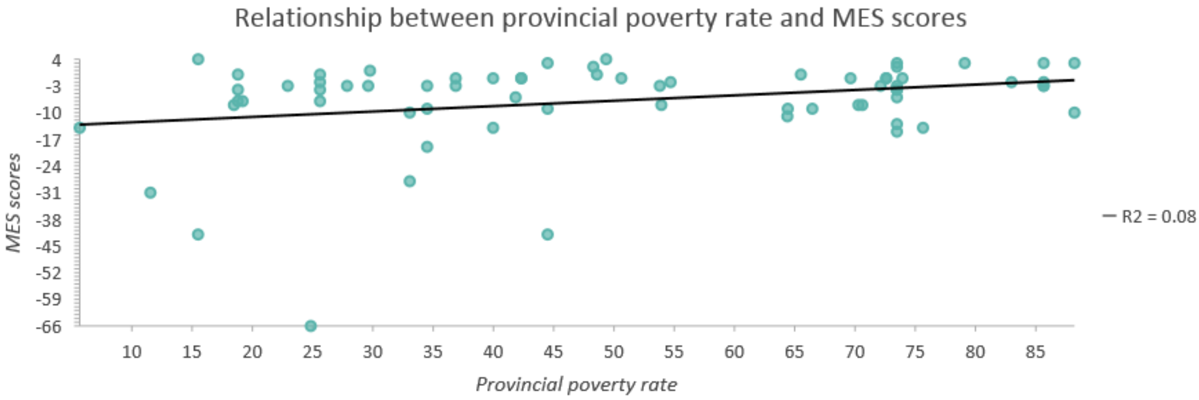

Figure 6: Scatterplot[4] of the relationship between provincial poverty rate and MES scores

[4] Slope and intercept of the trend line: y = -13.97653 + 0.14154 x

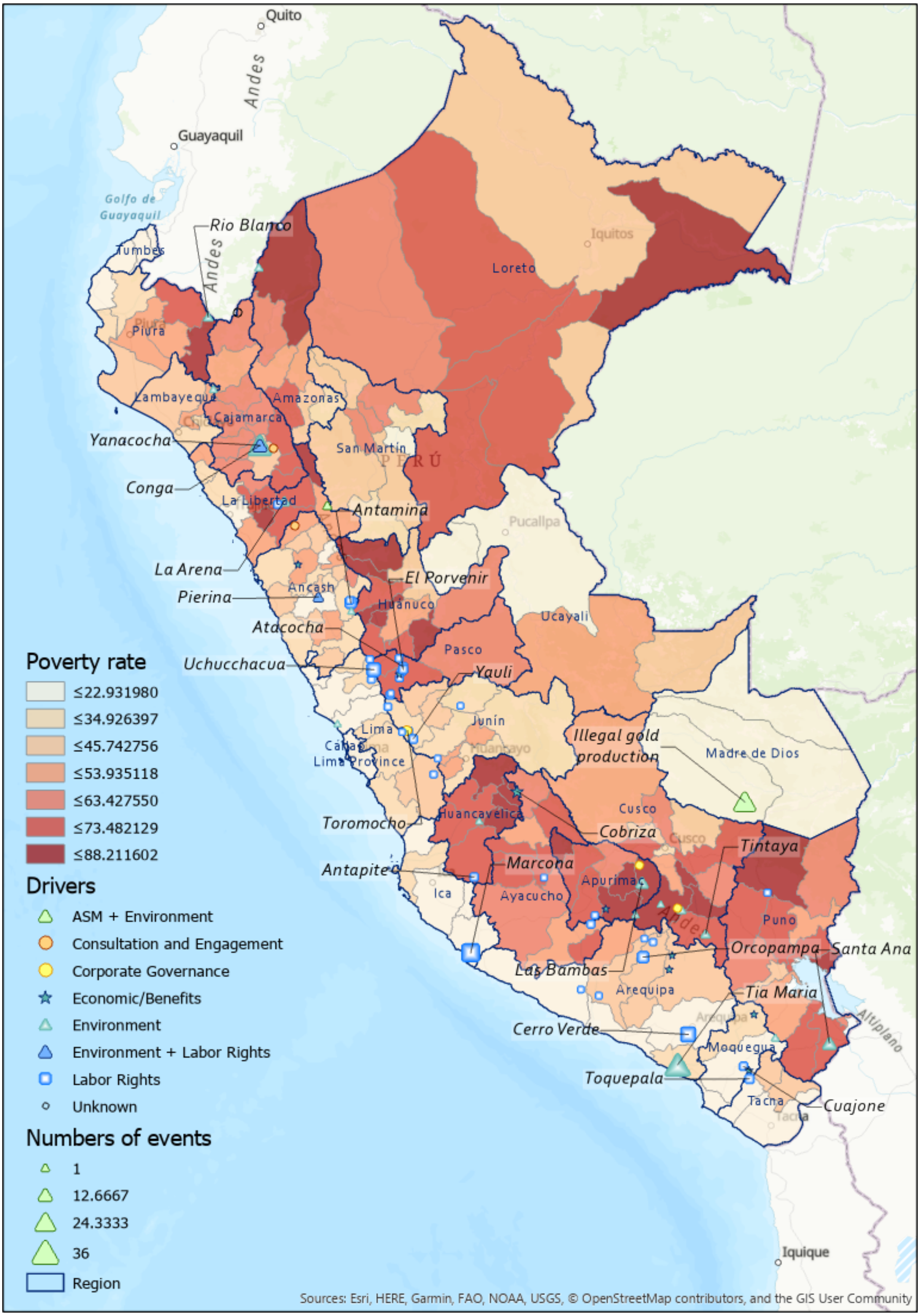

Figure 7: Map of provincial poverty in Peru (2009)

4) Elevation

Figure 8: Elevation map of Peru

As shown by Figure 34, the elevation in Peru ranges from -4 m to 6693m. Most mine sites are located in the Andes areas with relatively high elevation, except mine projects Marcona, Cerro Verde, Tia Maria and illegal gold production in Madre de Dios.

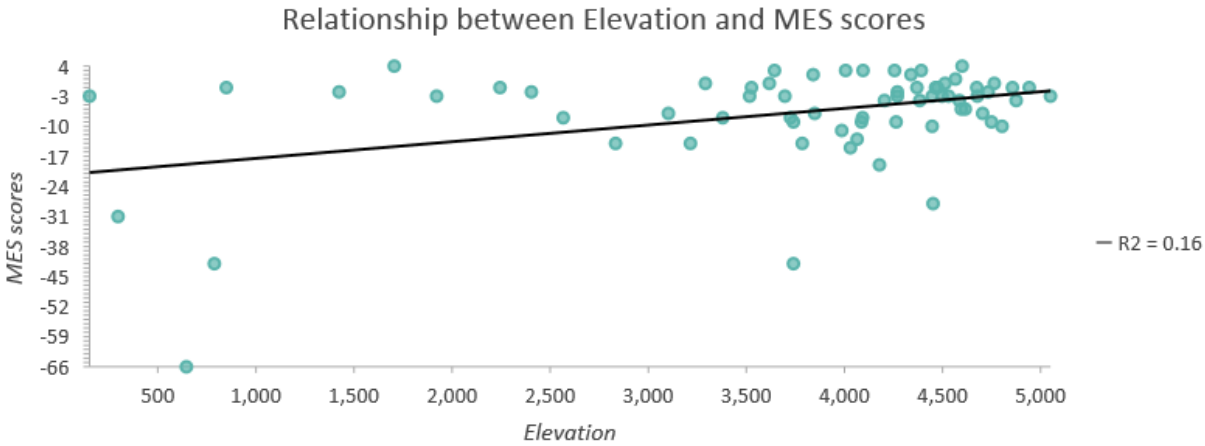

Figure 9: Scatterplot of the relationship between elevation and MES scores

The scatter plot and the trend line[5] indicate the positive relationship between altitude and MES scores. This confirm the hypothesis (H5) that the higher altitude a mine site is located, the less conflictive the mine site is, because traditional livelihoods in Peru such as agriculture and husbandry are limited by arid climates in high altitude regions, therefore locals are less resistance to mining.

5) Water supply

As shown by Figure 36, the annual precipitation of Peru ranges from 2 to 8053 mm/year. The annual rainfall in Switzerland is in the range of 812.8mm to 2387.6mm (Weather Atlas 2021). In comparison, Peru is also a country with ample water resources since the average value is around 1619.67mm/year. However, the geographical distribution is severely uneven. Most rainfall is in the Amazon rainforest in eastern Peru. Also, water bodies like rivers are mainly in the Amazon rainforest. The precipitation on the Pacific side is scarce. However, economic activity is also concentrated on the Pacific side of the Andes, with 94% of the country's extractive industry activities and two-thirds of agricultural production (INEI, 2010; as cited in Salem et al. 2018).

[5] Slope and intercept of the trend line: y = -21.37024 + 0.00388 x

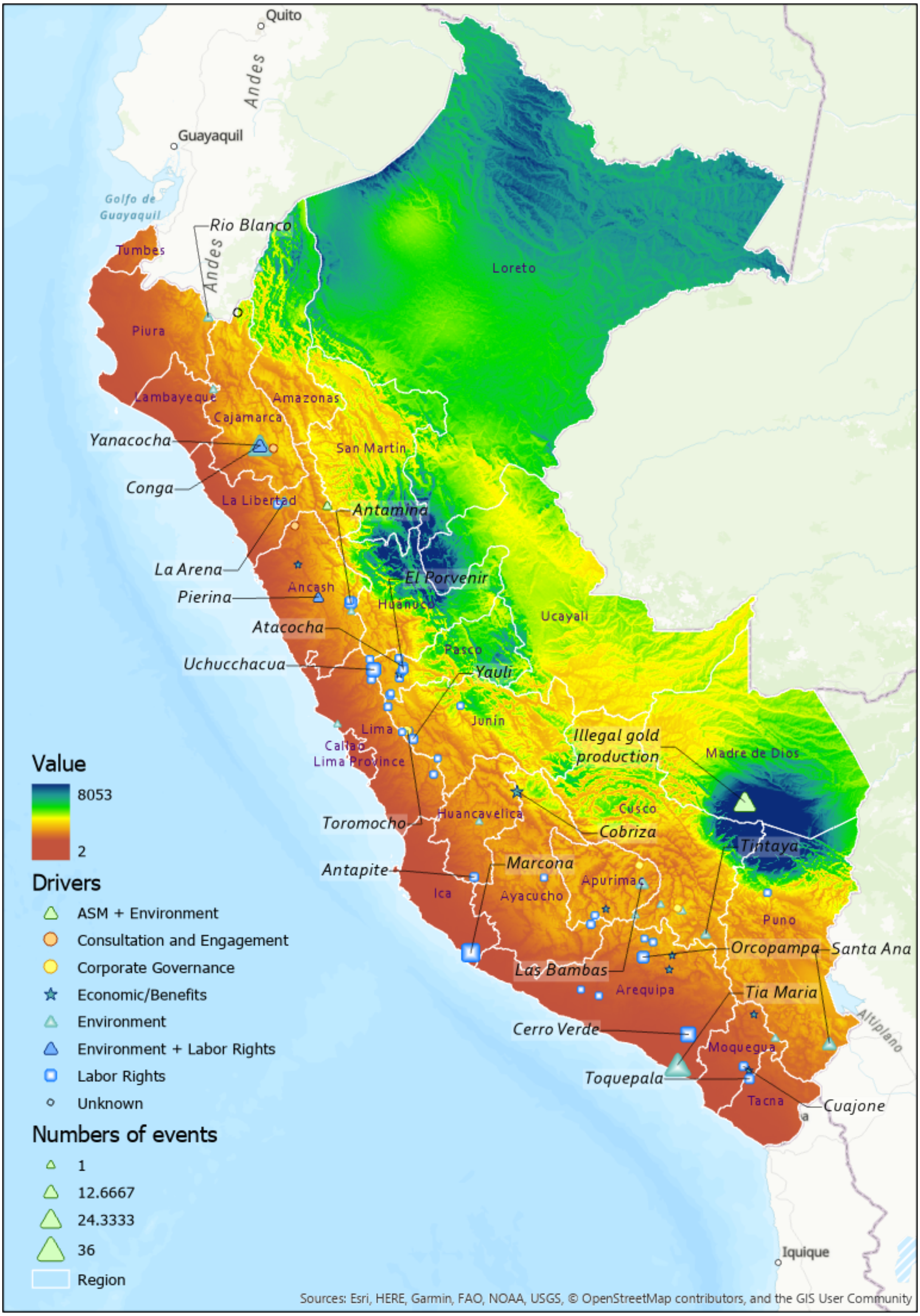

Figure 10: Map of annual precipitation in Peru

Since only seven mine projects have directly water-related issues such as insufficient water supplies and water pollution, according to the findings of the event dataset in the previous part. At this stage, there is no relationship between precipitation and MES scores due to the limited number of observations. Instead, a list of precipitation for these seven mine projects is provided (Table 15). All of them have precipitation values lower than the average of Peru's, but these don't indicate a water scarcity near these mine projects, except for Tia Maria. Therefore, at this stage, this hypothesis (H6) can not be confirmed.

Mine projects	Annual precipitation (mm)	MES scores
Tia Maria	15	-66
Conga	821	-42
Rio Blanco	1122	-14
Tintaya	779	-11
Santa Ana	801	-9
Parcoy	991	-8
Pierina	754	-7

Table 1: List of mine projects with water issues

The annual precipitation near the mine project Tia Maria is nearly 0. At the same time, the mine project is very closed to Tambo Valley, where intensive agricultural activities are located. The agriculture in Tambo Valley heavily relies on water resources from the Tambo river flowing through the valley. As shown by Figure 37, the areas colored in red next to Tia Maria are croplands in Tambo Valley. The closest community is district Cocachacra, among its 40000 residents, 2000 are plantation owners, 7000 are small holders and renters, and 8000 of them are day laborers (Dunlap 2019). Therefore, around half of the locals are involved in agricultural activities. This explained why locals protested violently over the years to resist the project.

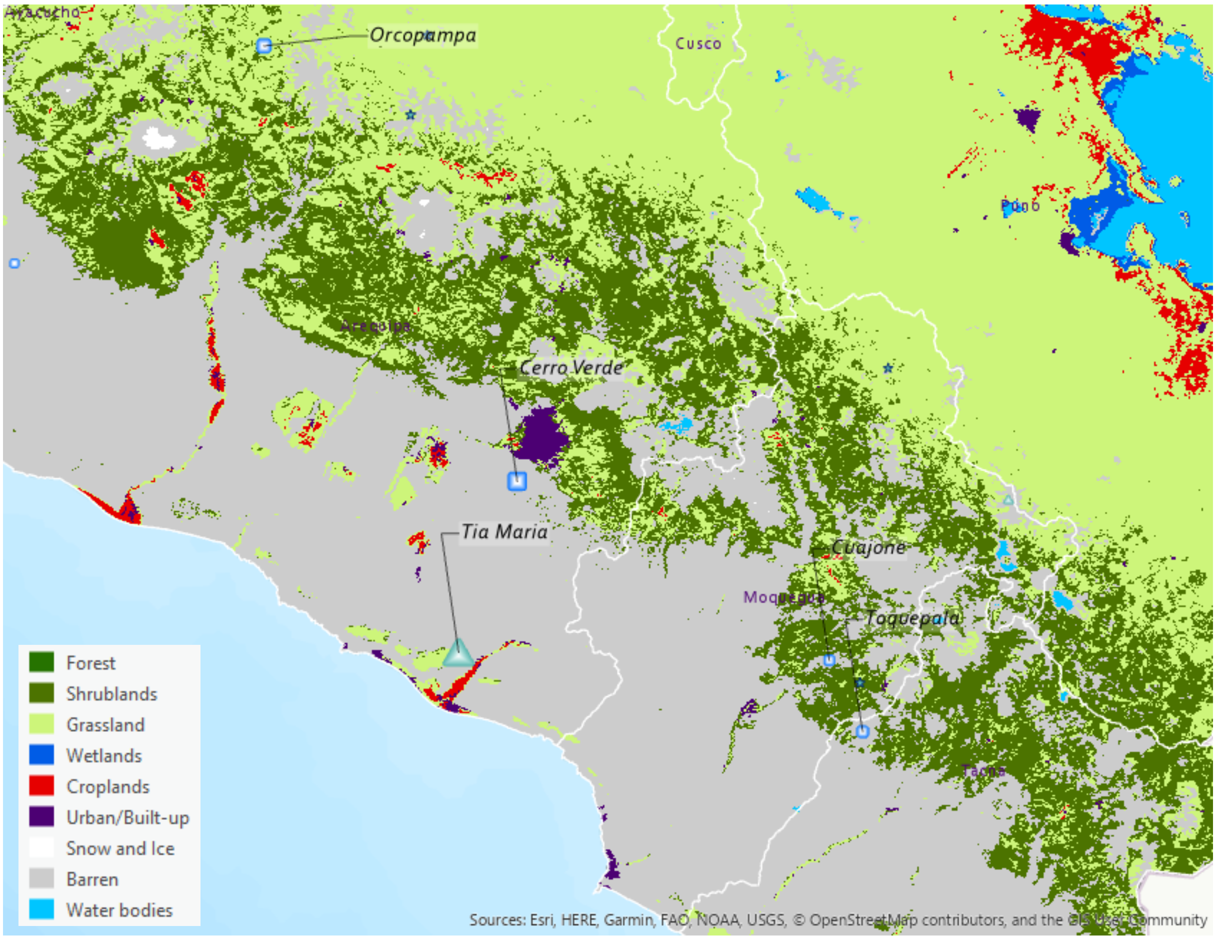

Figure 11: Landcover around Tia Maria

(MODIS/Terra+Aqua Land Cover Type, 2014)

The operator company Southern Copper has worked continuously to revise water supply plans but never resolve the conflicts. The project is still suspended, not to mention the economic loss due to the delay. From my perspective, such a solution can only solve one of the main problems behind the conflict. Main problems include not only environmental issues such as competition of limited water resources between mining and agriculture, but also social issues such as employment and the impact on traditional livelihoods. The job opportunities offered by Southern Copper are not only limited (hardly 600 workers) compared to those offered by agriculture (more than 20000 families) in the Valley, but they require technical expertise that will discriminate against the young, old and unskilled laborers (Dunlap 2019). The dimension of employment usually is not included in traditional CSR practices due to the difficulties of training mineworkers. This partially explains why the extensive CSR programs of Southern Copper are not helping Tia Maria get enough supports from the locals.

1.3 Limitations and outlook

Since the testing of hypotheses is mine project-based, due to the limited number of mine projects captured in the event dataset, the statistical significance of results at this stage is not available. Results are overviewing trends between MES and social, environmental, and economic characteristics around mine sites. Such results offer hypotheses for large-N analysis once the global dataset is complete.

Due to the limited number of recorded mine sites, each hypothesis was tested individually without considering the possible inter-correlation between different variables. For example, I assume that areas with less population density might overlay areas with high altitudes. In the future, with more observations in the event dataset, one can conduct map algebra analysis by giving weights to different raster layers (e.g., elevation, population, landcover), based on the direction and size of each coefficient calculated by regression analyses. The map algebra will produce a new raster layer that considers all variables after summing (or other math operations) cell values for all raster layers. Therefore, this method can produce a mining risk map for the location selection of mine sites in the future.

In addition, spatial data have uncertainties. Except for the DEM, which has a high spatial resolution of 30m, the spatial resolution of other raster layers is either 500m or 1000m. In the future, if one would like to conduct a spatial analysis regarding landcover, it might not be sufficient since some residential areas in the mountainous Andes are small and discrete, which might not be captured by satellite sensors. As for the method to extract values of layers for each point, I extract the values exactly where points are located. This might not be suitable in reality since we should consider the situation within a radius of, e.g., 10km or 25km around the mine site. For example, most mines are located in a barren land where population density is relatively low. However, there might be a large residential area 5km away. In comparison, the population density map only considers the count per 1 km2. So far, I can think of is to create, e.g., a 10km buffer around mine sites and calculating the characteristics within each buffer. However, the workload is high, and I didn't find an efficient way to implement this at this stage.