6th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Development of aggregated risk matrix for assessing the pipeline multiple natural disasters and their visualization using GIS

Aslan Babakhanov *[1]

[1]*Ministry of Science and Education Republic of Azerbaijan, Institute of Geography named after academician H.A. Aliyev, Azerbaijan*

Keywords
Risk assessment
Pipelines
Natural Disasters
Spatial Risk
Kernel Density Estimation

Abstract

This paper presents an approach for the computation and evaluating of an aggregated risk rating matrix to quantify the vulnerability of Baku-Tbilisi-Ceyhan pipeline to multiple natural disasters (ND) on the territory of Azerbaijan. The scope of (ND) encompasses earthquakes, floods, landslides, faults, mud volcanoes and soil erosions. Our proposed approach merges individual risk evaluations for each ND with certain intrinsic, but with generalized and aggregated calculation of final risk value smoothed by Kernel Density Estimation method. Further, we employ Geographic Information System (GIS) technology for the spatial representation of the computed risks across the pipeline network by using the final and aggregated risk values per special cell. This visual approach facilitates a better comprehension of spatially varied risks and supports the effective planning of risk mitigation measures. This visual approach aids in better understanding of the spatial distribution of risks, thereby supporting efficient strategizing of risk mitigation measures. The model was evaluated using the linear part of Baku-Tbilisi-Ceyhan pipeline as a case study, demonstrating the method's robustness and versatility. The methodology outlined in this study offers a rigorous and adaptable tool for ND risk assessment using GIS tools, but with some compute intensive methods. It aims to the enhancing pipeline safety and resilience on the territory of Azerbaijan

1. Introduction

Enhancing the resilience of the infrastructural components within the Azerbaijani hydrocarbon extraction complex to safeguard against potentially catastrophic effects of ND presents a critical and timely challenge in the scientific community. The pipeline systems, responsible for transferring hydrocarbon resources from the oil and gas deposits, constitute high-risk elements within this complex. Given the inherent flammability of these resources, they possess the capacity to initiate incidents with calamitous repercussions, resulting in irrecoverable ecological damage, loss of human lives, substantial financial and administrative losses, and infrastructural devastation (Lerche and Bagirov 2014).

As the role of pipeline interconnectivity expands among various countries, coupled with the rising incidents of natural disasters, the relevance and necessity of risk assessment methods are increasingly emphasized (Krausmann et al. 2011).

There are known ND on the territory of Azerbaijan that may directly have an impact to pipelines: earthquakes, heavy rainfalls and they are may trigger other events, such as landslides, mud volcanoes, floods, soil erosions and liquefaction (Amirova-Mammadova 2018).

It is common to visualize the impact and risk values as a single map per risk factor, but for bigger picture (here, using the single map) there is a need for an aggregation of risk values. Here comes the risk matrix, but with additional capabilities to evaluate the all required risk factor.

2. Method

In this study, the re-evaluated risk matrixes per weighted natural disaster will be used to assess each impact factor, smooth them using the Kernel Density Estimation (KDE) applied to a risk scores and at the end they will be mapped as a 2D array (Gramacki 2017). The

* **Corresponding Author**

*(aslan@babakhanov.az) ORCID ID 0000-0002-8790-7945

Cite this study

Babakhanov, A. (2023). Development of aggregated risk matrix for assessing the pipeline multiple natural disasters and their visualization using GIS. Intercontinental Geoinformation Days (IGD), 6, 121-125, Baku, Azerbaijan

values assigned to each matrix cells will be adopted in any of relational databases (Nasser 2014) or any convenient textual data format, like comma delimited text files, JSON files or any other data table format that can be used by widely used GIS tools like ESRI ArcGIS or QGIS tools.

2.1. Risk Matrix

To study the impact of natural disasters on pipelines, a risk matrix is an important component in representing the level of safety. For the risk calculation methods, I'll base it on a standard 5x5 risk matrix which evaluates hazards based on their likelihood (from rare to almost certain) and their consequences (from insignificant to catastrophic). Thus, five level is one of the optimal options for levels (Blokdyk 2018).

The more important is to identify the certain regions with certain hazards that may affect the pipelines (Han and Weng 2010). The fact that, region representing as a squared polygon area on the map, it could be of any size, from 10x10 meters up-to kilometers and classified by risk level. For this purpose, a distribution of risk levels must be prepared for squared area "Table 1".

This table is important part of all calculations and be included into the visualization of GIS maps.

Table 1. Risk level distribution per ND factor

Natural Disaster	Risk Level (1-5, 5 is highest)	Description	Color
Landslides	4	Soil movements in hilly or mountainous areas that can displace or damage pipelines	Yellow
Earthquakes	5	Ground shaking or rupture causing severe damages	Red
Flooding	3	Cause erosion or sedimentation	Yellow
Corrosion from soil chemistry	3	Certain soil conditions can enhance corrosion leading to pipeline material degradation	Yellow

2.2. Calculating Risk values

In order to deal with multiple events of hazards, each contribution to a single risk factor be aggregated to the singe risk score.

Firstly, calculate the individual risk scores ($\boldsymbol{R_i}$) for each hazard as per Equation 1.

$$\boldsymbol{R_i} = \boldsymbol{L_i} \times \boldsymbol{C_i} \qquad (1)$$

where $\boldsymbol{L_i}$ is the likelihood and $\boldsymbol{C_i}$ is the consequence of the $\boldsymbol{i^{th}}$ hazard. Once we have the individual risk scores for each hazard ($\boldsymbol{R_1, R_2, \ldots, R_n}$), then calculating the average risk score ($\boldsymbol{R_{avg}}$) as follows:

$$\boldsymbol{R_{avg}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{R_i} \qquad (2)$$

In this case, $\boldsymbol{n}$ is the total number of hazards or events being considered. Taking the average is just the one of many possible methods to aggregate the risk, but suitable for mapping purposes. Notably, the color coding of the risk scores would be: ***1-6*** as Green (low risk), ***7-12*** as Yellow (medium risk) and ***13-25*** as Red (high risk). Each value that falls into each score might be considered as a gradient color as well.

On the other hand, calculating an aggregate risk score using this method assumes each hazard is independent and equally significant, which may not always be the case. So, for such a complex situation, we will incorporate a weighting system to account for the varying significance of different hazards or correlations between events. Here we assign a weight ($\boldsymbol{w_i}$) to each hazard, which represents its relative importance or significance. This would typically be a value between 0 and 1, with total of all weights equal to 1. The next step is to calculate the weighed risk score ($\boldsymbol{R_{wi}}$) for each hazard:

$$\boldsymbol{R_{wi}} = \boldsymbol{w_i} \times \boldsymbol{R_i} \qquad (3)$$

then, sum all of the weighted risk scores to get the total risk score ($\boldsymbol{R_{total}}$):

$$\boldsymbol{R_{total}} = \sum_{i=1}^{n} \boldsymbol{R}_{\mathrm{wi}} \qquad (4)$$

and finally apply the color coding based on the risk score:

$$Color(\boldsymbol{R_{total}}) = \begin{cases} Green & if\ 1 \leq \boldsymbol{R_{total}} \leq 6 \\ Yellow & if\ 7 \leq \boldsymbol{R_{total}} \leq 12 \\ Red & if\ 13 \leq \boldsymbol{R_{total}} \leq 25 \end{cases}$$

This method gave is a risk score that takes into account both the individual risk scores and the relative importance of each hazard, but not considering the correlation between hazards.

2.3. Smoothing risk scores by categories

Mapping of granulated risk values over the map would give us coarse image even by dividing each risk score into sub-score. In order to smooth the risk values over the 2D map, we will use the Kernel Density Estimation (KDE) method to generate an estimation of the probability density function of multiple risk values (here small risk zones). So, while doing the KDE for each category, you will treat the risk values outside the category range as if they don't exist.

The equation (Equation 1) for each risk category (**c** as 1: *Green, 2: Yellow, 3: Red*) consider the data points ($\boldsymbol{x}_i, \boldsymbol{y}_i$) that fall within that category's risk score range.

$$\hat{f}_c(x,y) = \frac{1}{n_c h^2} \sum_{i=1}^{n_c} \frac{1}{2\pi} e^{-\frac{1}{2}\left[\left(\frac{x - x_i^{(c)}}{h}\right)^2 + \left(\frac{y - y_i^{(c)}}{h}\right)^2\right]}$$

So, risk levels (here categories) per impact, where $(\boldsymbol{n_c})$ is the number of data points in the $(\boldsymbol{i})^{th}$ category, $(\boldsymbol{h})$ is the bandwidth parameter, $(\boldsymbol{e})$ is the base of the natural logarithm and $x_i^{(c)}$ and $y_i^{(c)}$ are the n^{th} pair of 2D data points.

This equation (Equation 1) assumes, that the **X** and **Y** variables are independent for each category within the Gaussian kernel. The bandwidth parameter h is very important in KDE as it determines the smoothness of the density estimate. A small **h** will make the estimate very sensitive to the data (potentially overfit), while a large **h** will make the estimate very smooth (potentially underfit). It's often chosen using cross-validation or a rule-of-thumb method (Silverman 2018).

2.4. Processing

Before starting the analyses on certain area of interest (AIO), there were made an analysis of most evident natural disaster events occurred on a territory of Azerbaijan. That analyses were a key process to determine the minimal dimension for AIO, where de identified as square of 1 km. In fact, that size is not de-facto and can be changed from places, regions, cities and etc. Changing the size of AIO is adding the more analyses to the events and their effect. The risk level determination, risk weight and score calculation of AIO is show in Figure 1.

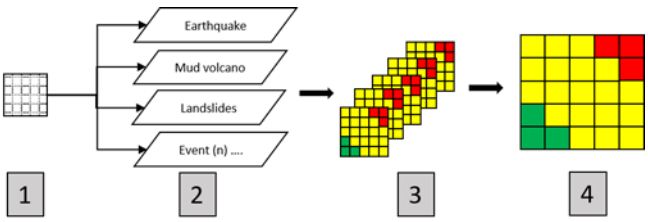

Figure 1. Processing of single AIO

The first step 1 is about to determine the AIO, where shape of pipeline must be at center of AIO. This will give us a proper and proportional GIS visualization of events area and pipeline linear part. At the step 2 we need identify the ND that may fit into this AIO. Surely, the dimension of AIO might be changed if any of identified ND has a broader effect. At the same step 2, we determine the risk parameters per event, their weights against the pipeline. Step 3 is about to calculation of risk matrixes per event. At the end (step 4) we aggregate the risks scores into a single risk score matrix.

2.5. Assessment challenges

One of challenges is the accurately description of the risk impact of a natural disaster that may affect the pipeline, then we need to disclose in a broader form the definition of the risk itself in order to give it a clearer vision of dimension of risk probability.

Here are the some of them (ordered by complexity)

1. Identification of risk parameters and their evaluations.
2. Determination of right risk weights per risk parameter.
3. Selection of proper level of granulation of area of interest to avoid coarsening.

The first challenge can be described when assessing the earthquakes effects on pipelines and triggered post events, including NATECH (Krausmann et al. 2016).

A three-parametric earthquake event that may affect the onshore pipeline is show in Figure 2. By adding color of risk category, we can see 3D view of earthquake effect on pipeline. This plot is showing only the risk at given location. Note, that location may change and values may vary.

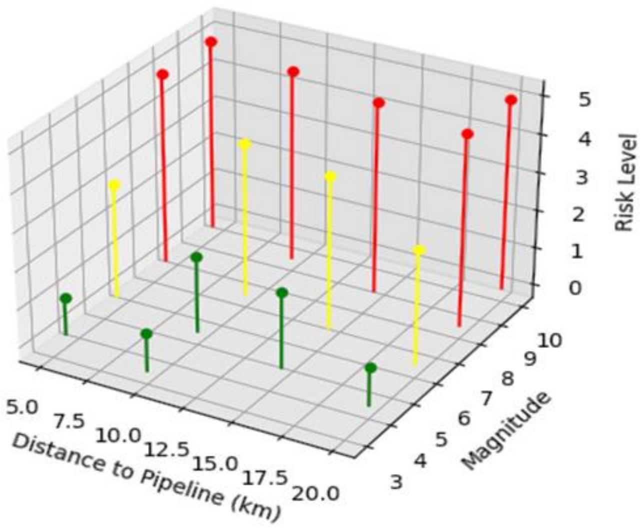

Figure 2. 3D earthquake risk assessment plot

The same 3D plot can be generated for other ND that may affect the pipelines, and they may have more parameters, adding assessment difficulties.

3. Results

As a result of study, we managed a good visualization of aggregated risk using GIS tools. Moreover, adding the buffered zones along the pipeline route providing the clearer picture of risk areas (Petersen 2020).

Figures from “Fig. 4” to “Fig. 7” are showing the evolution of aggregated fine-tuned risk scores from multiple natural hazards in a small region. The minimal area (which is AIO) is about 2 km² and whole assessed region is 100 km². The risk matrixes were based on mix of registered and assessed earth quakes, known faults and mud volcanoes.

Following tools were used to accomplish the visualizations:

- Maps were build using ESRI ArcMap 10.8
- Calculations were done in Python 3.8
- Output data format was CSV files

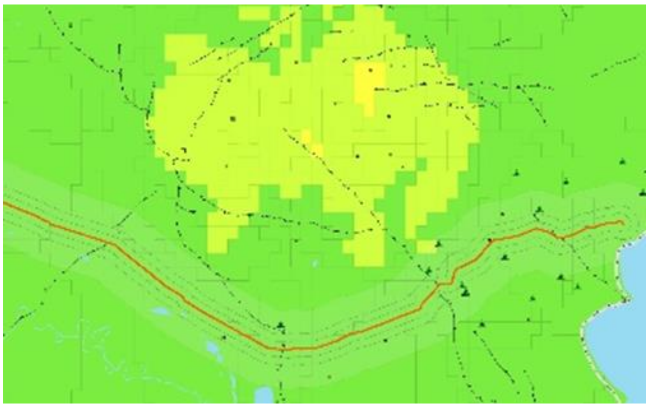

Figure 4. Coarse, aggregated risk scores

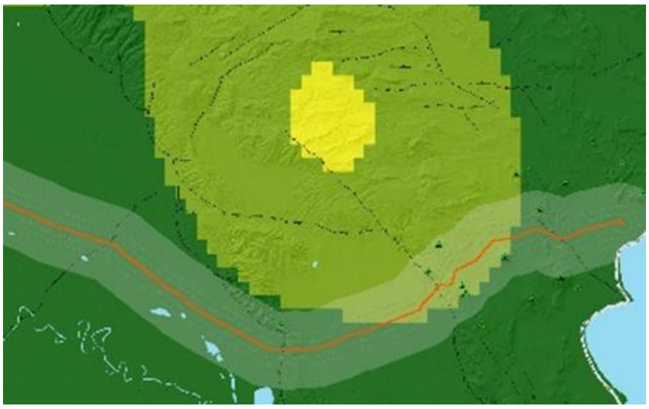

Figure 5. KDE based aggregated risk scores

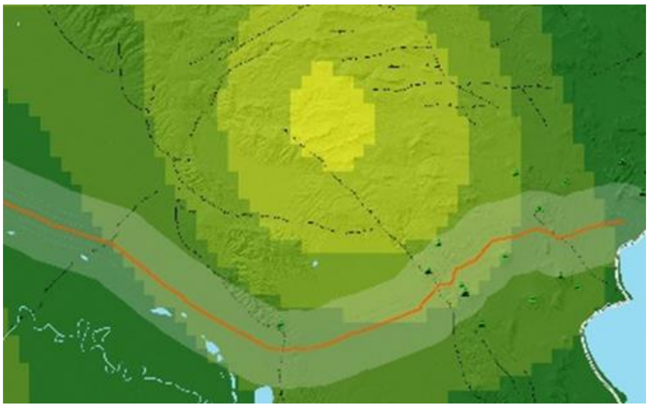

Figure 6. KDE based, small tuned risk scores

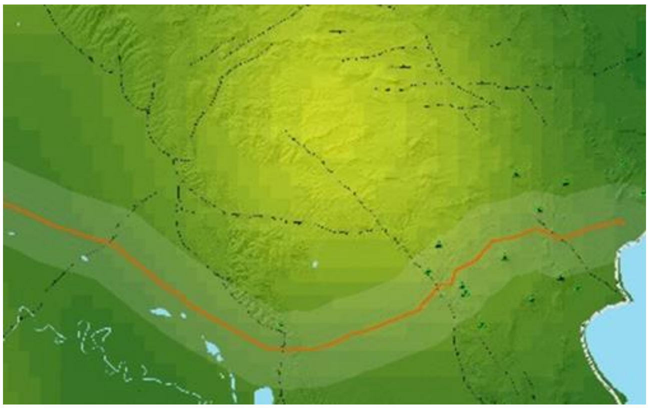

Figure 7. KDE based, fine-tuned risk scores

4. Conclusion

We identified several factors contributing to the difficulties of determination of risk weights:

1. **Complexity of Risk Interactions:** ND often do not occur in isolation. They can interact and compound in complex ways, making it difficult to accurately assign weights. For example, the risk posed by an earthquake could increase the risk of a subsequent landslide, but determining exactly how much additional risk this interaction creates can be challenging.
2. **Lack of Data:** For many risks, especially rare or unprecedented ones, there may be insufficient data to accurately estimate probabilities and impacts. This lack of data can lead to uncertainty in the risk weights.
3. **Over-simplification**: In an effort to simplify, there is a danger of oversimplifying the complexity of risks and their interactions by assigning a single risk weight. This might lead to an underestimation of the actual risk.
4. **Dependence on Model**: The assigned risk weights are only as good as the risk assessment model used. If the underlying model does not accurately represent the situation, the risk weights might be misleading.

Despite the difficulties, we were able to visualize the risk in aggregated form, using single map with fine-grained and tuned risk scores.

One of the best findings, is this model can be used to assess the other risks by using the same matrix representation and weight calculations and then we can implement the KDE method for fine tuning the final output if necessary. Also, the chained operations are giving us flexibility to include an additional function like risk score determination for relative AIO regions by implementing the Moore neighborhood (Ilachinski and Zane. 2001) algorithms.

Acknowledgement

This study was carried out as part of the master's thesis prepared by Aslan Babakhanov at Institute of Geography named after academician H.A. Aliyev.

References

Amirova-Mammadova, S. (2017). Pipeline Politics and Natural Gas Supply from Azerbaijan to Europe. Wiesbaden: Springer.

Blokdyk, G. (2018). Risk Matrix a Complete Guide. 5starcooks.

Gramacki, A. (2017). Nonparametric Kernel Density Estimation and Its Computational Aspects. Springer.

Han, Z., & Weng, W. (2010). An integrated quantitative risk analysis method for natural gas pipeline network. Journal of Loss Prevention in the Process Industries, 23(3), 428–436. https://doi.org/10.1016/j.jlp.2010.02.003

Ilachinski, A., & Zane. (2001). Cellular Automata – A Discrete Universe. Kybernetes, 32(4). https://doi.org/10.1108/k.2003.06732dae.007

Krausmann, E., Cruz, A. M., & Salzano, E. (2016). Natech Risk Assessment and Management: Reducing the Risk of Natural-Hazard Impact on Hazardous Installations. Elsevier.

Krausmann, E., Renni, E., Campedel, M., & Cozzani, V. (2011). Industrial accidents triggered by earthquakes, floods and lightning: lessons learned from a database analysis. Natural Hazards, 59(1), 285–300. https://doi.org/10.1007/s11069-011-9754-3

Lerche, I., & Bagirov, E. (2014). Impact of Natural Hazards on Oil and Gas Extraction. Springer.

Nasser, H. (2014). Learning ArcGIS Geodatabases. Packt Publishing Ltd.

Petersen, K. (2020). Visualizing Risk: Drawing Together and Pushing Apart with Sociotechnical Practices.

Silverman, B. W. (2018). Density Estimation for Statistics and Data Analysis. Routledge.