Introduction to Geoprocessing and Python in GIS

Chenyi Weng

SSCI586 - GIS Programming and Customization

Professor Jennifer N. Swift

September 8th, 2024

1. Introduction

In this project, my goal was to assess the proximity of schools in Los Angeles County to earthquake fault lines and determine the potential risk levels for these schools. By using ArcGIS Pro's geoprocessing tools, this project built a workflow to identify schools near fault lines and classify them based on their risk due to proximity. However, this analysis helps identify which schools may require seismic retrofitting or other precautionary measures.

2. Study Area

Los Angeles County is an area prone to seismic activity due to its proximity to the San Andreas Fault and various other fault lines. It includes a mixture of urban, and rural environments, making it an interesting area to examine for earthquake preparedness, particularly in public and private schools. Figure 1 represents a map outlining the study area, including boundaries of Los Angeles County and the distribution of earthquake faults.

Figure1. Boundary of Study Area

3. Data Inputs and Preparation

The data used for this project includes public and private school locations; also, both datasets were initially in the WGS 1984 projection, which was maintained consistently throughout the analysis. The earthquake fault data consisted of multiple fault classes, representing varying levels of seismic hazard potential. To focus the analysis on our specific area of interest, the original datasets were clipped to the boundaries of Los Angeles County, as shown in Figure 2. This step was important to refine the scope and concentrate the analysis on the relevant geographic region.

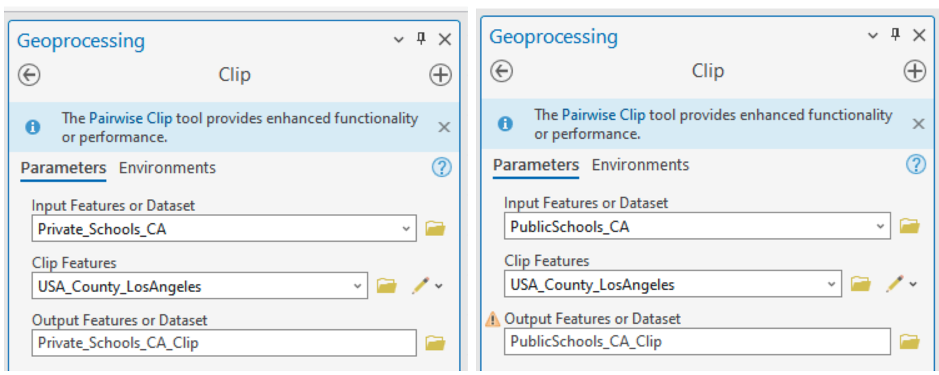


Figure 2. Clipping Public and Private Schools to Los Angeles County Boundary

4. Method

4.1 Model Builder Workflow

The workflow was built using the ModelBuilder tool in ArcGIS Pro, and the main steps are as follows: First, the public schools, private schools, and earthquake fault datasets were clipped to retain only the data within Los Angeles County. Next, the private and public school datasets were merged to create a unified school dataset. Then, a 1-mile buffer was applied around the fault lines to identify schools within a 1-mile proximity to the faults. In the meanwhile, to categorize the schools based on proximity to the faults, I added a new field called RiskLevel and used the Calculate Field tool to assign risk levels (Class A, B, or C) as shown in Figure 3; also, the result is visualized in the ModelBuilder workflow as shown as Figure 4.

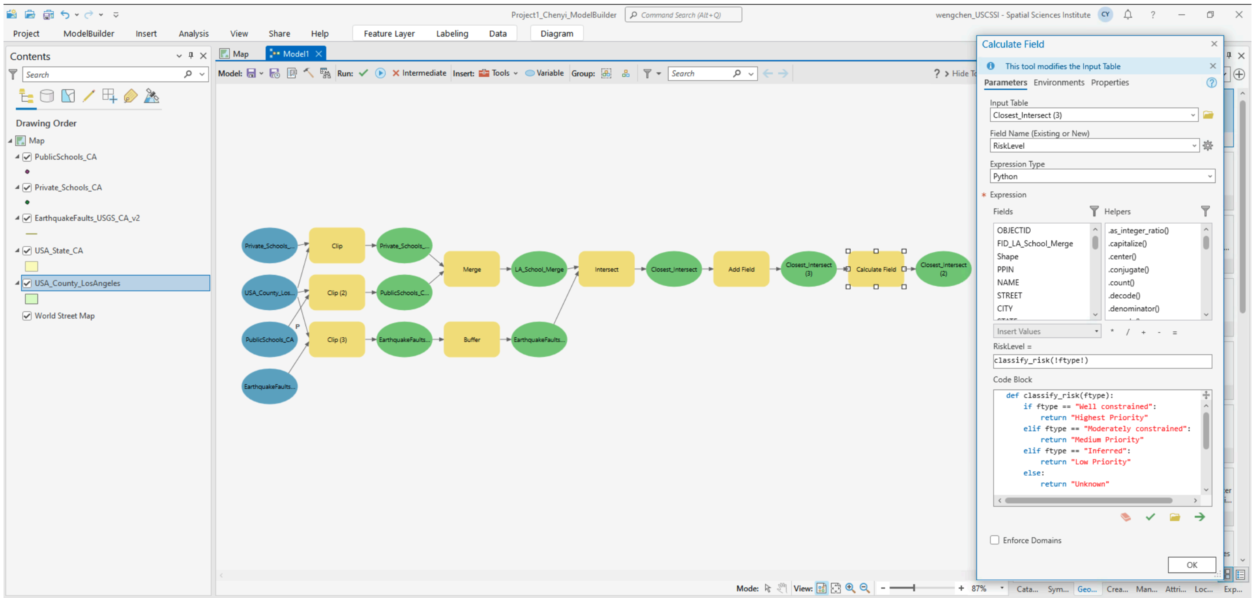


Figure 3. Used the Calculate Field Tool.

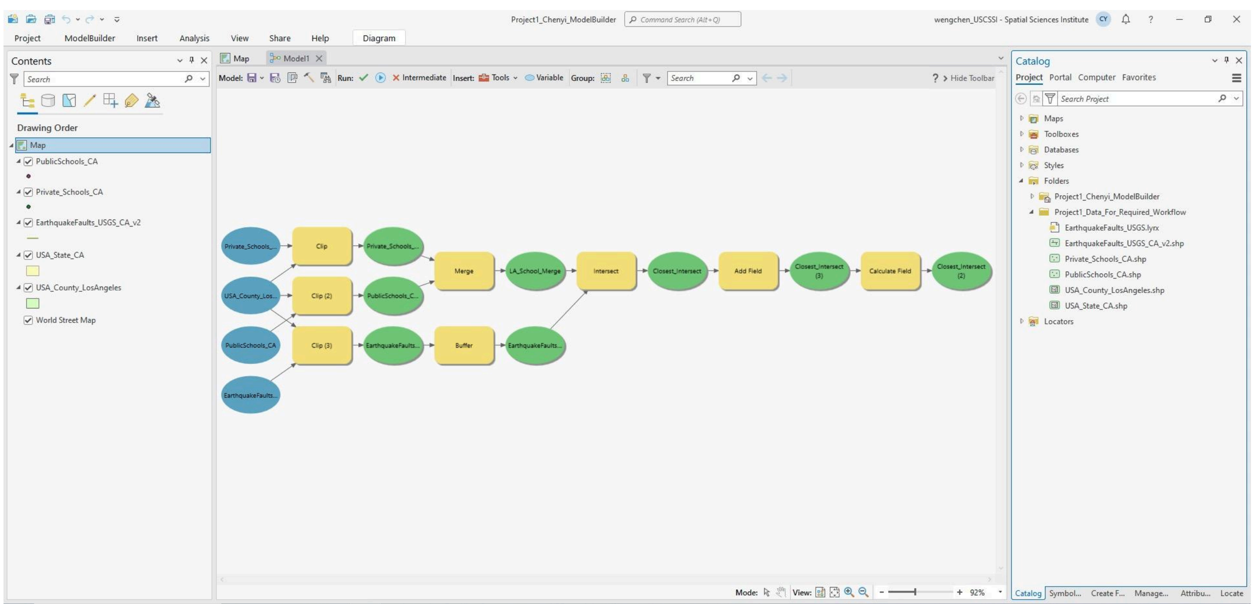
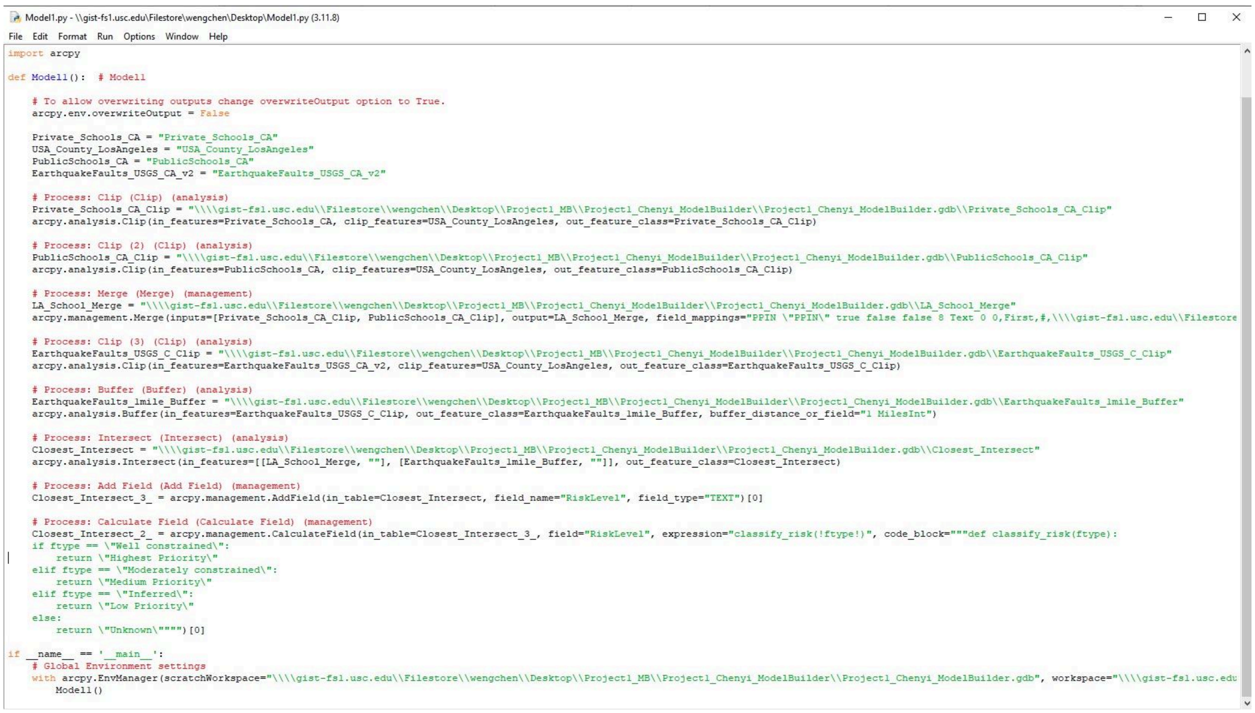


Figure 4. Workflow Designed by ModelBuilder

4.2 Risk Calculation

After intersecting the buffered fault lines with the school data, a new field was added to classify the schools based on their proximity to faults. Schools were categorized into high, medium, or low priority for safety measures, depending on how close they were to the fault lines. This classification is implemented using a Python script embedded in the ModelBuilder workflow, as shown in Figure 5.



Model1.py - \\gist-fs1.usc.edu\Filestore\wengchen\Desktop\Model1.py (3.11.8)

File Edit Format Run Options Window Help

```python
import arcpy

def Model1():  # Model1

    # To allow overwriting outputs change overwriteOutput option to True.
    arcpy.env.overwriteOutput = False

    Private_Schools_CA = "Private_Schools_CA"
    USA_County_LosAngeles = "USA_County_LosAngeles"
    PublicSchools_CA = "PublicSchools_CA"
    EarthquakeFaults_USGS_CA_v2 = "EarthquakeFaults_USGS_CA_v2"

    # Process: Clip (Clip) (analysis)
    Private_Schools_CA_Clip = "\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb\\Private_Schools_CA_Clip"
    arcpy.analysis.Clip(in_features=Private_Schools_CA, clip_features=USA_County_LosAngeles, out_feature_class=Private_Schools_CA_Clip)

    # Process: Clip (2) (Clip) (analysis)
    PublicSchools_CA_Clip = "\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb\\PublicSchools_CA_Clip"
    arcpy.analysis.Clip(in_features=PublicSchools_CA, clip_features=USA_County_LosAngeles, out_feature_class=PublicSchools_CA_Clip)

    # Process: Merge (Merge) (management)
    LA_School_Merge = "\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb\\LA_School_Merge"
    arcpy.management.Merge(inputs=[Private_Schools_CA_Clip, PublicSchools_CA_Clip], output=LA_School_Merge, field_mappings="PPIN \"PPIN\" true false false 8 Text 0 0,First,#,\\\\gist-fs1.usc.edu\\Filestore

    # Process: Clip (3) (Clip) (analysis)
    EarthquakeFaults_USGS_C_Clip = "\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb\\EarthquakeFaults_USGS_C_Clip"
    arcpy.analysis.Clip(in_features=EarthquakeFaults_USGS_CA_v2, clip_features=USA_County_LosAngeles, out_feature_class=EarthquakeFaults_USGS_C_Clip)

    # Process: Buffer (Buffer) (analysis)
    EarthquakeFaults_1mile_Buffer = "\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb\\EarthquakeFaults_1mile_Buffer"
    arcpy.analysis.Buffer(in_features=EarthquakeFaults_USGS_C_Clip, out_feature_class=EarthquakeFaults_1mile_Buffer, buffer_distance_or_field="1 MilesInt")

    # Process: Intersect (Intersect) (analysis)
    Closest_Intersect = "\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb\\Closest_Intersect"
    arcpy.analysis.Intersect(in_features=[[LA_School_Merge, ""], [EarthquakeFaults_1mile_Buffer, ""]], out_feature_class=Closest_Intersect)

    # Process: Add Field (Add Field) (management)
    Closest_Intersect_3_ = arcpy.management.AddField(in_table=Closest_Intersect, field_name="RiskLevel", field_type="TEXT")[0]

    # Process: Calculate Field (Calculate Field) (management)
    Closest_Intersect_2_ = arcpy.management.CalculateField(in_table=Closest_Intersect_3_, field="RiskLevel", expression="classify_risk(!ftype!)", code_block="""def classify_risk(ftype):
    if ftype == \"Well constrained\":
        return \"Highest Priority\"
    elif ftype == \"Moderately constrained\":
        return \"Medium Priority\"
    elif ftype == \"Inferred\":
        return \"Low Priority\"
    else:
        return \"Unknown\"""")[0]

if __name__ == '__main__':
    # Global Environment settings
    with arcpy.EnvManager(scratchWorkspace="\\\\gist-fs1.usc.edu\\Filestore\\wengchen\\Desktop\\Project1_MB\\Project1_Chenyi_ModelBuilder\\Project1_Chenyi_ModelBuilder.gdb", workspace="\\\\gist-fs1.usc.edu
        Model1()
```

Figure 5. Risk Classification Python Script

The calculated field assigned a risk level to each school, which was visualized in the final map as shown in Figure 6. Therefore, schools classified as highest priority (red), medium priority (yellow), and low priority (green) are displayed, offering a clear geographic representation of which areas are most at risk.

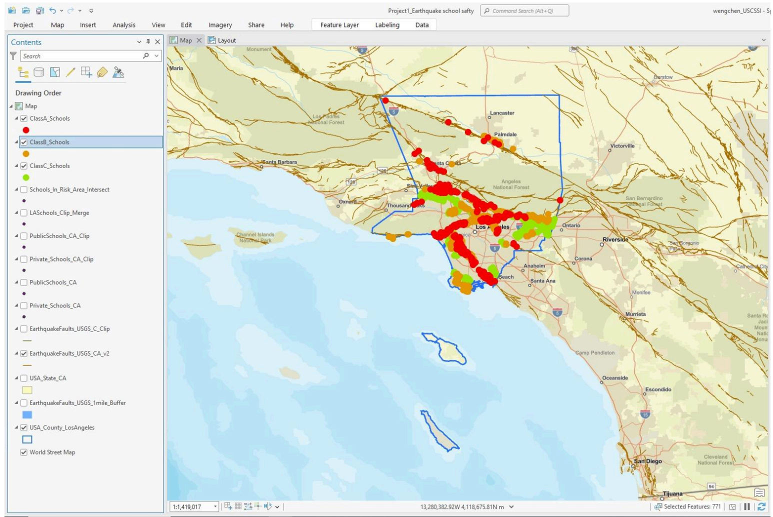


Figure 6. Risk Level Map of Schools in Los Angeles County

5. Results Visualization

As shown in Figure 7, the final map provides a comprehensive view of the schools in relation to earthquake faults. The color coding based on risk level allows for easy identification of schools that require immediate attention in terms of earthquake preparedness.

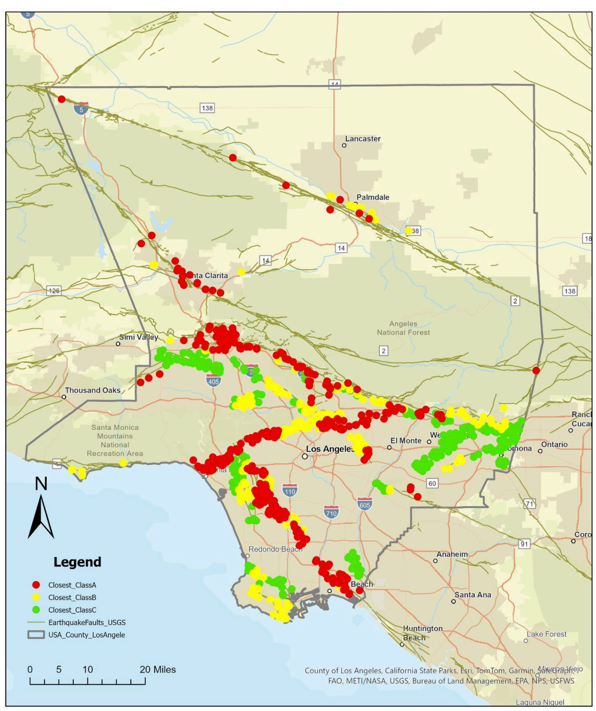
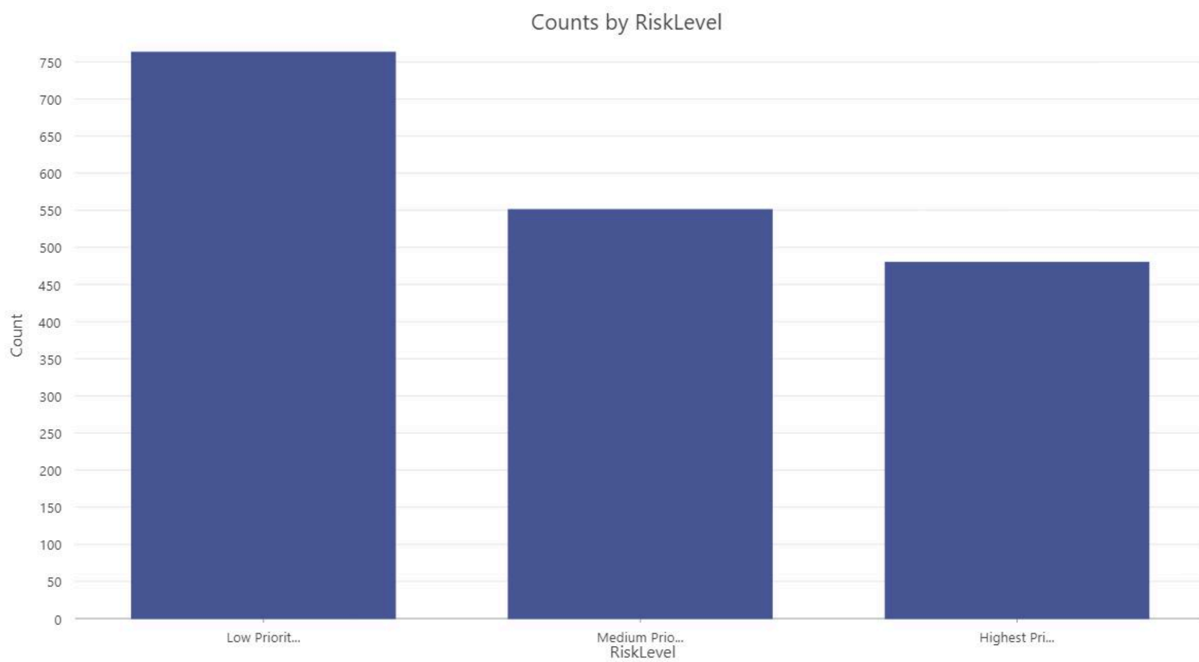


Figure 7. Risk Level Map of Schools in Los Angeles County

Additionally, a summary chart of the number of schools in each risk category is presented in Figure 8. This bar chart illustrates that the majority of schools fall into the low priority category, while a significant number of schools are in the high priority category, indicating a substantial risk that needs to be addressed.



RiskLevel	Count
Low Priority	763
Medium Priority	551
Highest Priority	480

Figure 8. Count of Schools by Risk Level

6. Discussion

There were several challenges faced during the completion of this project. One limitation was the simplification of earthquake risk assessment using proximity alone. In real-world scenarios, factors like the building age, construction type, and retrofitting status of schools also play an important role in seismic vulnerability, as highlighted in Leon et al. (2007). For example, while some schools may be close to a fault, they may have undergone seismic retrofitting that significantly reduces their vulnerability. Conversely, schools further from a fault may still be highly vulnerable if built before modern seismic building codes were enforced. Additionally, the buffer distance of 1 mile was arbitrarily selected based on available research and may not fully capture the risk variation across the fault types. Lastly, debugging the ModelBuilder workflow when converting it into Python was a challenge, especially in handling file paths and correctly exporting geoprocessing results. This process, however, allowed me to better understand Python's role in automating GIS processes.

Reference

Leon, Lorraine A, Shari A Christofferson, James F Dolan, John H Shaw, and Thomas L Pratt. "Earthquake-by-Earthquake Fold Growth above the Puente Hills Blind Thrust Fault, Los Angeles, California: Implications for Fold Kinematics and Seismic Hazard." *Journal of Geophysical Research - Solid Earth* 112, no. B3 (2007): B03S03-n/a. https://doi.org/10.1029/2006JB004461.