Technical Appendix for ERP Dissertation

Author: Ruoxin Zhang

1. Overview

This appendix provides detailed technical steps to ensure the reproducibility of the dissertation research. It covers data acquisition, QGIS preprocessing, Python-based analysis, and verification of outputs.

2. Data Sources

Dataset	Source	Format	Resolution/Unit
OpenStreetMap (POIs)	Geofabrik	.gpkg	Points; lines
Global Relative Deprivation Index (GRDI)	NASA SEDAC (CIESIN)	.tif	Raster, 1 km
WorldPop	WorldPop Project	.tif	Raster, 100m
GADM Level-2 boundaries	GADM	.shp	Polygons

3. QGIS Preprocessing (using Uganda as example)

3.1 Import OSM Data

Source: Geofabrik extracts for Ghana and Uganda.

File format: gis_osm_pois_free_1.shp; gis_osm_transport_free_1.shp

3.2 Select Infrastructure Categories

Filter by fclass attribute:

(a) From gis_osm_pois_free_1.shp

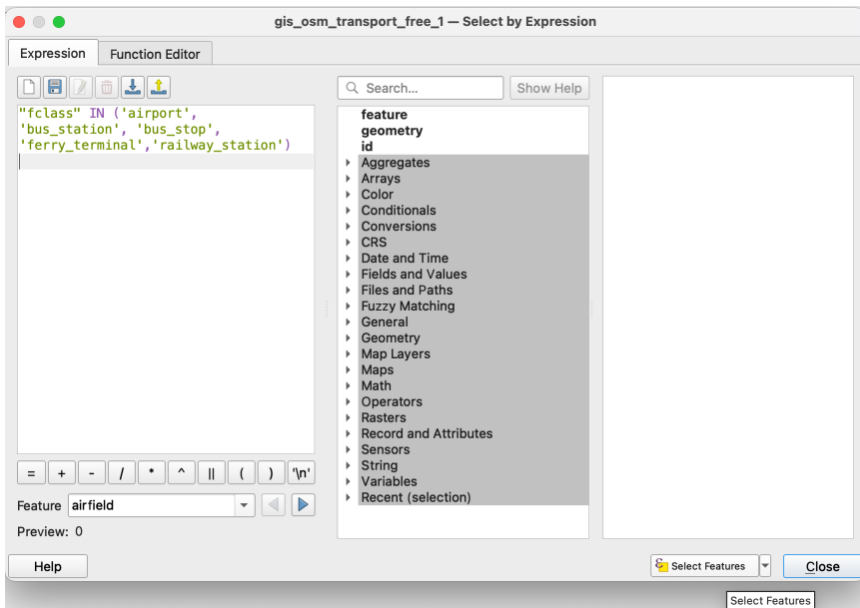- Education → school, college, university, kindergarten, education institutions

- Health Care → hospital, clinic, doctors, pharmacy

- Commerce → supermarket, convenience, market_place

- Green Space → park, playground, pitch, playground, sport_centre
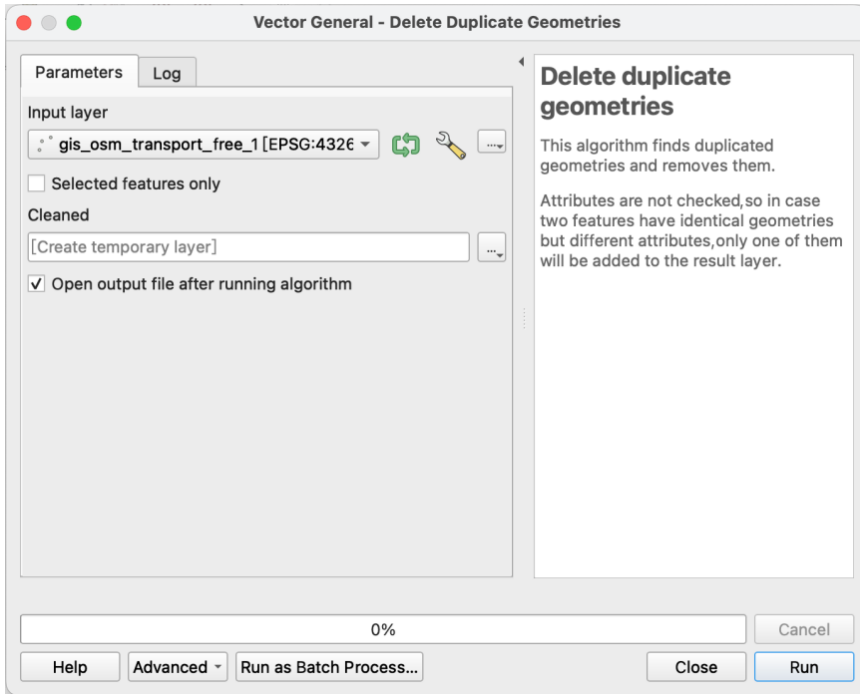
(b) From gis_osm_transport_free_1.shp

- Public Transport → airport, bus_station, bus_stop, ferry_terminal, and railway_station

3.3 Geometry Operations

(a) Delete Duplicate Geometries

  - Menu: Processing → Toolbox → Geometry General → Delete Duplicate Geometries

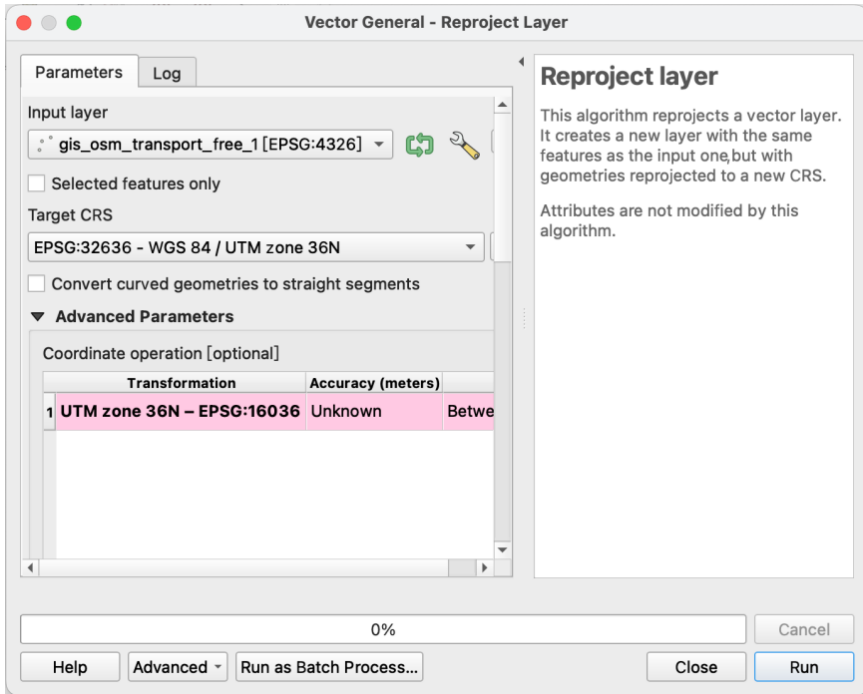 - Purpose: remove overlapping or duplicate POIs to avoid double-counting.

(b) Reproject Layers

Ghana → EPSG:32630 (UTM Zone 30N)

Uganda → EPSG:32636 (UTM Zone 36N)

Menu: Processing → Toolbox → Geometry General → Reproject Layer

3.4 Export to GeoPackage

Export each category as .gpkg

Naming convention: <theme>_reprojected.gpkg

Example: shop_reprojected.gpkg, education_reprojected.gpkg

4. Python Analysis Workflow

RUN the Code in the Github with Jupter Notebook

4.1 Accessibility Indicators

4.2 Composite Infrastructure Index (PCA)

4.3 Regression Models (OLS & SEM)

4.4 Verification

Outputs included in /outputs/: composite index (CSV), regression results (CSV).

Minor differences may occur if OSM or WorldPop data are re-downloaded.

5. Dependencies

Python 3.10

QGIS 3.42

Packages: geopandas, shapely, rasterio, rasterstats, scikit-learn, statsmodels, pysal, matplotlib