

Introduction

The spatial feature data set consists of three parts, namely POI data, land use data and road network data. The spatial feature dataset is retrieved from OpenStreetMap (OSM), which is an open source world map and provides a variety of geographic features. Three kinds of features that influence the emission and diffusion of PM_{2.5} particles are selected, including POI, road network and land use.

1. POI Data

Description:

This part of data is stored in “Spatial_Data\POI”, denoting the information of POI. In order to facilitate the data statistics of the paper, we integrated POI data and road network data into one file (“Spatial_Data\POI-Road.xlsx”). The POI data is in columns 2 through 9 in the file.

POIs are of point type. For each grid, we count the number of POIs that belongs to different categories shown in Table 1 and regard each category as a feature.

TABLE 1 POI CATEGORY

| Category | POIs |
|---------------------|---|
| Vehicle | gasstation, car wash, parking and underground parking |
| School | university, school, kindergarten, college |
| Transportation spot | traffic signal, roundabout, crossing, motorway junction and turning circle |
| Shopping | supermarket, shopping mall, department store and convenience store |
| Food and beverage | restaurant, fast food, pub, cafe and bar |
| Entertainment | theater, nightclub, cinema, park and sports center |
| Transport | railway station, railway halt, tram stop, bus stop, bus station and taxi rank |
| Hotel | hotel, motel and guesthouse |

Schema:

Each row in POI-Road.xlsx stands for the number of POI features in each grid, and the columns, separated by comma, are defined as follows:

| ID | Vehicle | School | Transporta- tion spot | Shopping | Food and beverage | Entertai- ment | Transport | Hotel |
|----|---------|--------|--------------------------|----------|----------------------|-------------------|-----------|-------|
|----|---------|--------|--------------------------|----------|----------------------|-------------------|-----------|-------|

The ID is represented by a number from 1 to 2500, representing 2500 grids.

2. Road Network Data

Description:

This part of data is stored in “Spatial_Data\Road”, denoting the information of road network. In order to facilitate the data statistics of the paper, we integrated POI

data and road network data into one file(“Spatial_Data\POI-Road.xlsx”). Among them, road network data are in columns 10 to 14 of the document.

Road features comprise motorway, trunk, primary, secondary and tertiary roads, which are line features. We record the total length of each road feature in each grid.

Schema:

Each row in POI-Road.xlsx stands for the Length of road network features in each grid, and the columns, separated by comma, are defined as follows:

| ID | Motorway | Primary | Secondary | Tertiary | Trunk |
|----|----------|---------|-----------|----------|-------|
|----|----------|---------|-----------|----------|-------|

The ID is represented by a number from 1 to 2500, representing 2500 grids.

3. Land Use Data

Description:

This part of data is stored in “Spatial_Data\Landuse”, denoting the information of land use.

Land use features includes forest, park, residential, industrial and commercial areas, which are polygon features. In order to facilitate the data statistics of the paper, we put five key features of land use into onedocument(“Spatial_Data\Landuse-data.xlsx”).

Schema:

Each row in Landuse-data.xlsx stands for the area of land use features in each grid, and the columns, separated by comma, are defined as follows:

| ID | Fclass | Area | Percentage | Landuse_id |
|----|--------|------|------------|------------|
|----|--------|------|------------|------------|

The ID is represented by a number from 1 to 2500, representing 2500 grids.

The Fclass is represented by 5 features of land use.

The area represents the size of each feature occupying each grid.

Percentage is the percentage of each feature that occupies the area of each grid.

Landuse_id is represented by 5 numbers, such as business is represented by the number 1;The forest is represented by the number 2.