Spatial Disaggregation of Population Data with 3D Building Information

—— A case study of Deqing County

FAN Deqing, ZHAO Xuesheng, Qiu Yue

1China Univ. of Mining & Technology (Beijing)

2018-11-20

Main Contents

1 Introduction
2 Methodology
3 Result & analysis
4 Applications
5 Conclusions
1 Introduction

- **Why?**
  - statistical information on socio-economic activities is widely available,
  - aggregated to country or regional administrative units,
  - useful for assessments,
  - smooth out spatial variations in impact

- **Definition** — Spatial disaggregation are processes by which information at a coarse spatial scale is translated to finer scales, while maintaining consistency with the original dataset [Monteiro et al. 2018].

- **Objective** — Provide more localized estimates and spatial analysis.
An example of Deqing

<table>
<thead>
<tr>
<th>Name (town)</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wukang</td>
<td>89944</td>
</tr>
<tr>
<td>Fuxi</td>
<td>26008</td>
</tr>
<tr>
<td>Xiazhuhu</td>
<td>23999</td>
</tr>
<tr>
<td>Wuyang</td>
<td>52180</td>
</tr>
<tr>
<td>Luoshe</td>
<td>20553</td>
</tr>
<tr>
<td>Zhongguan</td>
<td>43856</td>
</tr>
<tr>
<td>Moganshan</td>
<td>31643</td>
</tr>
<tr>
<td>Qianyuan</td>
<td>49644</td>
</tr>
<tr>
<td>Leidian</td>
<td>37592</td>
</tr>
<tr>
<td>Xinan</td>
<td>31730</td>
</tr>
<tr>
<td>Xinshi</td>
<td>72395</td>
</tr>
<tr>
<td>Yuyue</td>
<td>33297</td>
</tr>
</tbody>
</table>

Disaggregation without Geospatial Information

In early stage, as lack of auxiliary data related to population, **Negative Index Model** [Feng & Zhou 2003; Wu & Gao 2010], **Nuclear Density Estimation Model** [Lu et al. 2002; Yan et al. 2011], etc were often used.

- **Principle** – (urban geography) population density decreases from the city center to the periphery.
- **Advantages** - simple model, simulation of continuous population distribution; suitable for large and medium-sized cities population density simulation.
- **Insufficiency** - The value of city center and bandwidth $\tau$ is subjective, not suitable for small cities and rural areas.
Spatial information used for disaggregation

Recently, various types and resolutions of population-related auxiliary data can be obtained, such as:

- land cover,
- traffic network
- DEM,
- water system,
- night lighting,
- OSM,
- mobile phones,
- .......

Many population data disaggregation methods have been developed, which can be divided into 4 categories:

Existing methods for disaggregation with spatial information

- Dasymetric mapping method
- Regression method
- Multi-factor synthesis method
- Spatio-temporal simulation method
A comparative analysis of existing disaggregation methods

<table>
<thead>
<tr>
<th>Types</th>
<th>Related factors</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Dasymetric mapping</td>
<td>Population, types of land cover, Topography, traffic network, impervious surface, etc.</td>
<td>Model simple &amp; easy, ensures the total population unchanged, suitable for fine-scale population spatialization.</td>
<td>Difficult accurate determine the weight of population allocation in each sub-area.</td>
</tr>
<tr>
<td>Bi-d [Holt et al. 2004; Langford 2007]; Tri-d [Mennis 2003; Lloyd 2016]; Multi-d [Su et al. 2010]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Regression method</td>
<td>The area of all types of land, and corrected by DEM, residential spots, night lighting, OSM data.</td>
<td>Model needs fewer parameters, is easy to model, results are controllable, suitable for large scale population spatialization.</td>
<td>Difficult to reveal the difference of population distribution under the same land type, and limited by the problem of light pixel overflow.</td>
</tr>
<tr>
<td>[Zhuo et al. 2005; Gallego et al. 2011; Malone 2012; Lu et al. 2013; Rosina et al. 2017]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Multi-factors synthesis</td>
<td>Population, land use, water factors, transportation network, River system, DEM, city size and location, residential areas, etc.</td>
<td>Comprehensive considering the influence of natural, economic and social factors, The results of the model are convincing.</td>
<td>The fusion weight is more subjective, and the index is changeable, which increases the complexity and redundancy of the model.</td>
</tr>
<tr>
<td>[Dobson et al 2000; Liu et al. 2003; Yue et al. 2003; Liao et al. 2010; Yao et al 2017; Monteiro 2018]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Spatio-temporal simulation</td>
<td>Demographic data, mobile location data (i.e. cell phone), etc.</td>
<td>suitable for describing spatial dynamic distribution of population in urban areas, and can estimate the permanent population</td>
<td>Poor results in country rural &amp; poor areas.</td>
</tr>
<tr>
<td>[Deville et al. 2014; Bakillah 2014; Lwin et al. 2016; Chen et al. 2018]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Strategy for disaggregation in this study

- **According to:**
  - the area distribution of Deqing County,
  - High-resolution land-cover data,
  - population statistics (in towns),

- **Dasymetric & area weighting method** will be adopted.
2 Procedure used in this study

a) Dasymetric - Dividing into residential areas and non-residential areas.

b) Area weighting - The residential areas should be weighted according to 6 types of residence.

c) Population calculation -

d) Spatial rasterilation – according to 30m×30m cell.
# Classification scheme of building density

By the density and height of buildings in residential areas, it will be divided into 6 types [according to the "Survey Contents and Indicators of Geographical Conditions (No. GDPJ 01—2013)"]:

<table>
<thead>
<tr>
<th>Types</th>
<th>description</th>
<th>Building density</th>
<th>Number of floors</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-M</td>
<td>High density &amp; Multi-floor building</td>
<td>( \geq 50% )</td>
<td>( \geq 4 )</td>
</tr>
<tr>
<td>L-M</td>
<td>Low density &amp; Multi-floor building</td>
<td>&lt; 50%</td>
<td>( \geq 4 )</td>
</tr>
<tr>
<td>H-L</td>
<td>High density &amp; Low-floor building</td>
<td>( \geq 50% )</td>
<td>&lt; 4</td>
</tr>
<tr>
<td>L-L</td>
<td>Low density &amp; Low-floor building</td>
<td>&lt; 50%</td>
<td>&lt; 4</td>
</tr>
<tr>
<td>M-S</td>
<td>Multi-floors &amp; Single building</td>
<td>( \geq 4 )</td>
<td></td>
</tr>
<tr>
<td>L-S</td>
<td>Low-floors &amp; Single building</td>
<td>&lt; 4</td>
<td></td>
</tr>
</tbody>
</table>

**Example -1**

[H-M]

[L-M]
Distribution map of six types

Types of building distribution

Legend
- L-M
- H-M
- L-S
- H-L

Weight determination for disaggregation

- The weight $p$ of a resident cell is
  \[ p = \lambda \times h \]
  \[
  \lambda - \text{building density in a resident cell;}
  
  h - \text{the average of all building floors in a cell.}
  \]

- The population $n$ of a resident cell is
  \[
  n_i = \frac{S_i P_i}{\sum_{i=1}^{m} S_i P_i} N
  \]
  \[
  S_i - \text{area of a resident cell;}
  
  N - \text{the whole population number in an administrative unit.}
  \]
Main Contents

1. Introduction
2. Methodology
3. Result & analysis
4. Applications
5. Conclusions

Spatial distribution of disaggregated population

Spatial distribution of population in Deqing County
Overly map - rural area

Sample points verification & error analysis
Main Contents

1 Introduction
2 Methodology
3 Result & analysis
4 Applications
5 Conclusions

Quantitative assessment of SDGs indicators

- Indicator 3.8.1- coverage the basic health services;
- Indicator 4.a.1- allocation of educational resources;
- Indicator 9.1.1- urban traffic
  - a. The proportion of rural population living within 2 km of the whole season highway;
  - b. Traffic accessibility;
  - c. X hour life circle
SDGs— indicator 3.8.1  Coverage of basic health services

Deqing County has:
- general hospitals - 3
- township hospitals - 19
- Health service stations - 134

Accessibility of general hospitals

Distribution frequency and cumulative frequency of service population of general hospitals
Accessibility of township hospitals

Access to township hospitals in Deqing County

<table>
<thead>
<tr>
<th>Time Range</th>
<th>Distribution Frequency (%)</th>
<th>Cumulative Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>53.277</td>
<td>53.277</td>
</tr>
<tr>
<td>5-10</td>
<td>39.164</td>
<td>92.441</td>
</tr>
<tr>
<td>10-15</td>
<td>6.670</td>
<td>99.110</td>
</tr>
<tr>
<td>15-20</td>
<td>0.812</td>
<td>99.922</td>
</tr>
<tr>
<td>20-25</td>
<td>0.077</td>
<td>99.998</td>
</tr>
<tr>
<td>25-30</td>
<td>0.002</td>
<td>100.000</td>
</tr>
<tr>
<td>&gt;30</td>
<td>0.000</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Distribution frequency and cumulative frequency of service population of township hospitals

Accessibility of health service stations

Access to health service stations in Deqing County

<table>
<thead>
<tr>
<th>Time Range</th>
<th>Distribution Frequency (%)</th>
<th>Cumulative Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>92.689</td>
<td>92.689</td>
</tr>
<tr>
<td>5-10</td>
<td>7.146</td>
<td>100</td>
</tr>
<tr>
<td>10-15</td>
<td>0.165</td>
<td>100</td>
</tr>
<tr>
<td>15-20</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>20-25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>&gt;25</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Distribution frequency and cumulative frequency of service population in health service station
At the end of 2017, Deqing County had:

- 17 primary schools
- 21 junior middle schools and 1 special school.
- 5 senior secondary schools.

### Distribution map of schools in Deqing County

#### Legend
- primary school
- junior high school
- high school

### Distribution of school bus

#### Legend
- elementary schools
- school bus sites
- highways
- national roads
- provincial roads
Accessibility of primary schools in Deqing County

Distribution frequency and cumulative frequency of service population in primary schools

Accessibility of junior high schools in Deqing County

Distribution frequency and cumulative frequency of service population in junior high schools
Accessibility of senior high schools in Deqing County

Distribution frequency and cumulative frequency of service population of general high schools

Appl. -Indicator 9.1.1

- SDGs-Indicator 9.1.1-
  - The proportion of rural population living within 2 km of the whole season highway;
  - Traffic accessibility;
  - X hour life circle

Tab. Proportion of population from X km to the road

<table>
<thead>
<tr>
<th>Indicator name</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proportion of rural population living within 500 meters</td>
<td>99.997%</td>
<td>99.997%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>The proportion of rural population living within 1000 meters</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>The proportion of rural population living within 2000 meters</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Indicator 9.1.1a- population living within 2 km

SDGs-Indicator 9.1.1a- The proportion of rural population living within 2 km of the whole season highway;

500m seasons road buffer and population distribution of Deqing county in 2016

9.1.1b- Traffic accessibility to urban

SDGs-Indicator 9.1.1- b. Traffic accessibility;

Map of time required for each location to reach county in Deqing county
9.1.1c- X hour life circle to town

SDGs-Indicator 9.1.1-c) X hour life circle

Map of The time required for each location to reach the nearest town in Deqing county

<table>
<thead>
<tr>
<th>Distribution Frequency (%)</th>
<th>0-5</th>
<th>5-10</th>
<th>10-15</th>
<th>15-20</th>
<th>20-25</th>
<th>25-30</th>
<th>30-35</th>
<th>35-40</th>
<th>&gt;40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11.78</td>
<td>51.13</td>
<td>31.22</td>
<td>4.27</td>
<td>0.67</td>
<td>0.68</td>
<td>0.24</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>Cumulative Frequency (%)</td>
<td>11.78</td>
<td>62.91</td>
<td>94.13</td>
<td>98.40</td>
<td>99.07</td>
<td>99.75</td>
<td>99.99</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Main Contents

1. Introduction
2. Methodology
3. Result & analysis
4. Applications
5. Conclusions
5 Conclusions

Populations, aggregated to a town administrative units in this study, is disaggregated to 30m×30m cell by spatial information. The main objective is to assess some SDGs indicators with a fine quantitative mode and spatial analysis.

- Dasymetric area weighting method is used in this research, in which Dasymetric mapping is by high resolution land-cover data and area weighting by 3D building information.
- Through the sample (50 villages) validation, the average accuracy is about 77.4%.
- As a case study of Deqing, quantitatively assess some SDGs indicators, such as health care (3.8.1), education(4.a.1), urban traffic (9.1.1), and accessibility analysis.

Thanks

China University of Mining & Technology (beijing)