COVID-19 Epidemic Spatio-Temporal Dataset User Guide

School of Geography and Information Engineering

China University of Geosciences (Wuhan)

February 18, 2020

1. Introduction

The **COVID-19 Epidemic Spatio-Temporal Dataset** intends to provide a multi-scale dataset of the novel coronavirus (COVID-19) epidemic data from December 31, 2019 to the end of the epidemic around the global. This dataset aims at providing authoritative, open and multi-scale COVID-19 epidemic information to scientists worldwide for providing suggestions for the epidemic.

The COVID-19 epidemic dataset includes statistical data at three scales: global, provincial (only available for China), and prefectural (only China). Statistical data at each scale contains six text files, i.e. new confirmed cases, new recovered cases, new death cases, total confirmed cases, total recovered cases, and total death cases.

2. Data Sources

COVID-19 global epidemic data (outside of China) comes from People's Daily (http://society.people.com.cn/GB/369130/431577/index.html, one of Chinese top ten authoritative Internet platforms). COVID-19 Chinese epidemic data at provincial and prefecture-level cities are gathered from the National, Provincial and Autonomous Region's Health Commission (short as the Health Commission). This dataset is planned to be weekly updated.

3. Data Format

The epidemic data is recorded in text file (i.e. .txt format) (shown in Table 1). The first row stands for the name of each attribute, including the Chinese and English name of the region and the date recording in T-yyyy-mm-dd format (e.g. February 20, 2020 was recorded as T20200204). The following rows records the value of each attribute field. For example, if 65 new recovered cases were found in Wuhan on February 4th, 2020 and 0 recovered in Xianning on February 4th, 2020, then it should be mark as 65 in the T20200204 column for Wuhan and 0 for Xianning. "-1" in the dataset indicates that the exact number of confirmed, recovered or death cases in this

area was not published by the Health Commission on this day.

Table 1 Data Format——Take the newly recovered cases as an example

Region_CN	Region_EN	T20200204	T20200205	T20200206	T20200207	...
武汉市	Wuhan	65	63	103	164	...
咸宁市	Xianning	0	1	3	6	...
...	...	...	...	...	...	...

4. Applications

Firstly, this dataset can be adapted in visualizing temporal variations of the epidemic under global, provincial (China) and prefecture (China) scales (shown in Figure 1).

Overlaying a vector base map, temporal-spatial changes of the epidemic under the scales of global, provincial (China) and prefecture (China) (shown in Figures 2-4) can be visualized. Vector data of countries can be obtained from https://gadm.org/download_world.html, and the vector data of Chinese provinces and prefectures can be obtained from http://www.webmap.cn/commres.do?method=dataDownload.

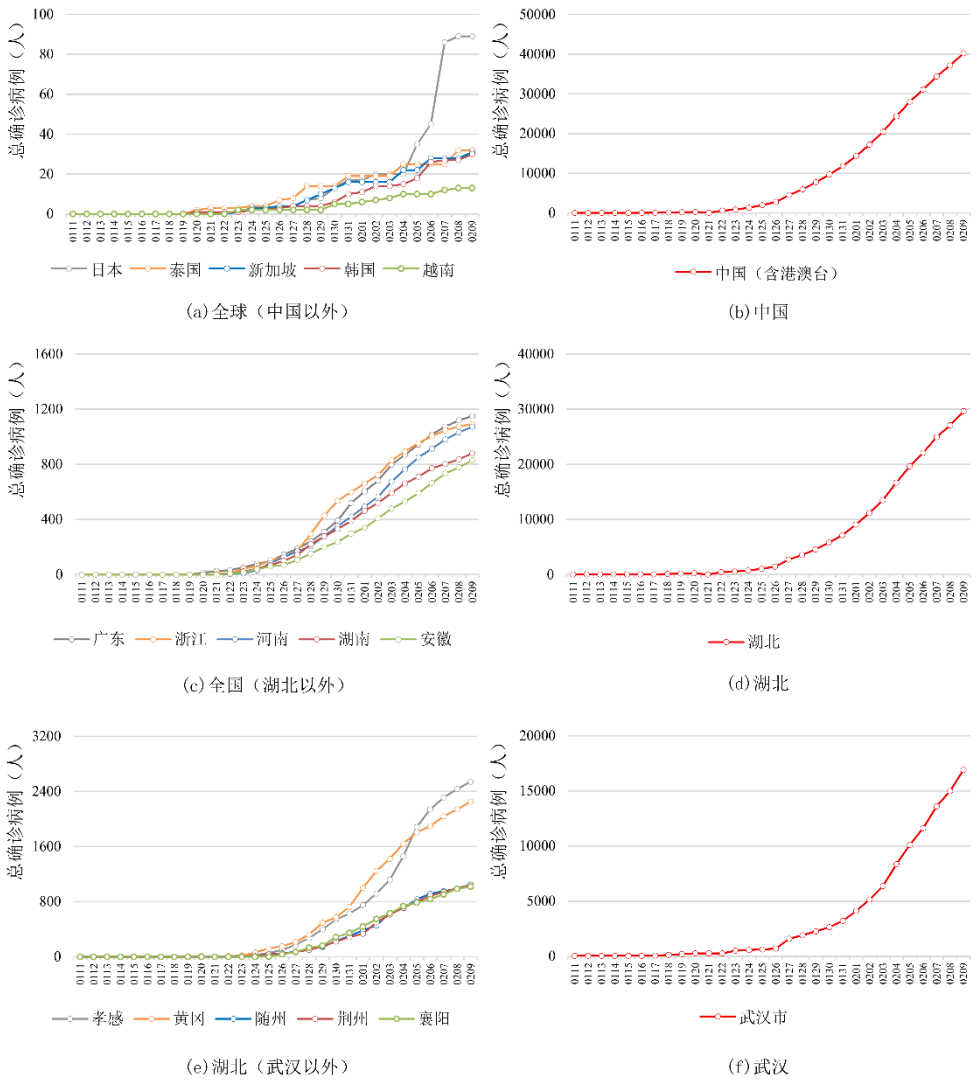

Figure 1 Temporal variation diagram based on COVID-19 epidemic data

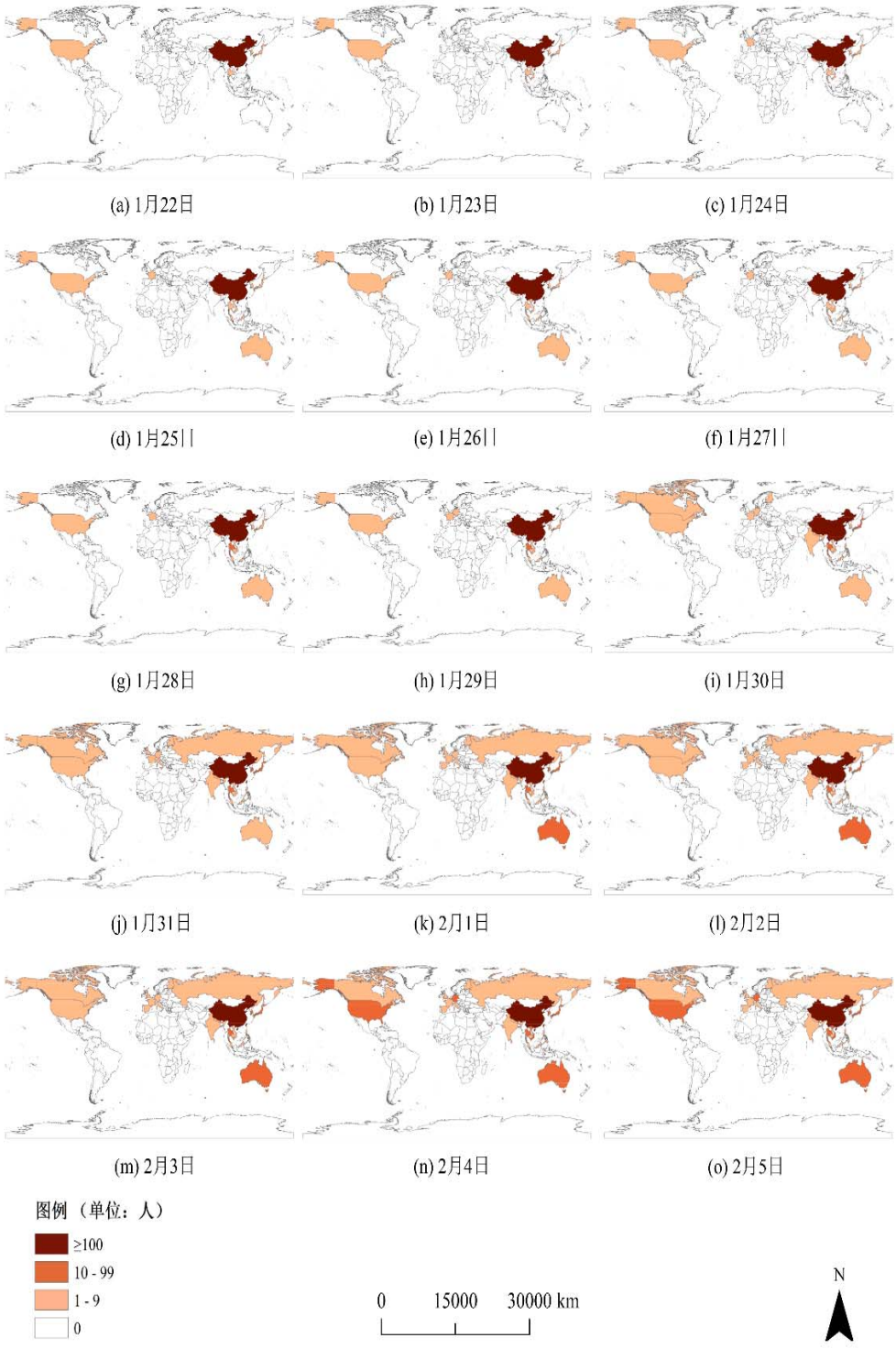

Figure 2 Temporal-spatial evolution map based on COVID-19 epidemic data (global scale)

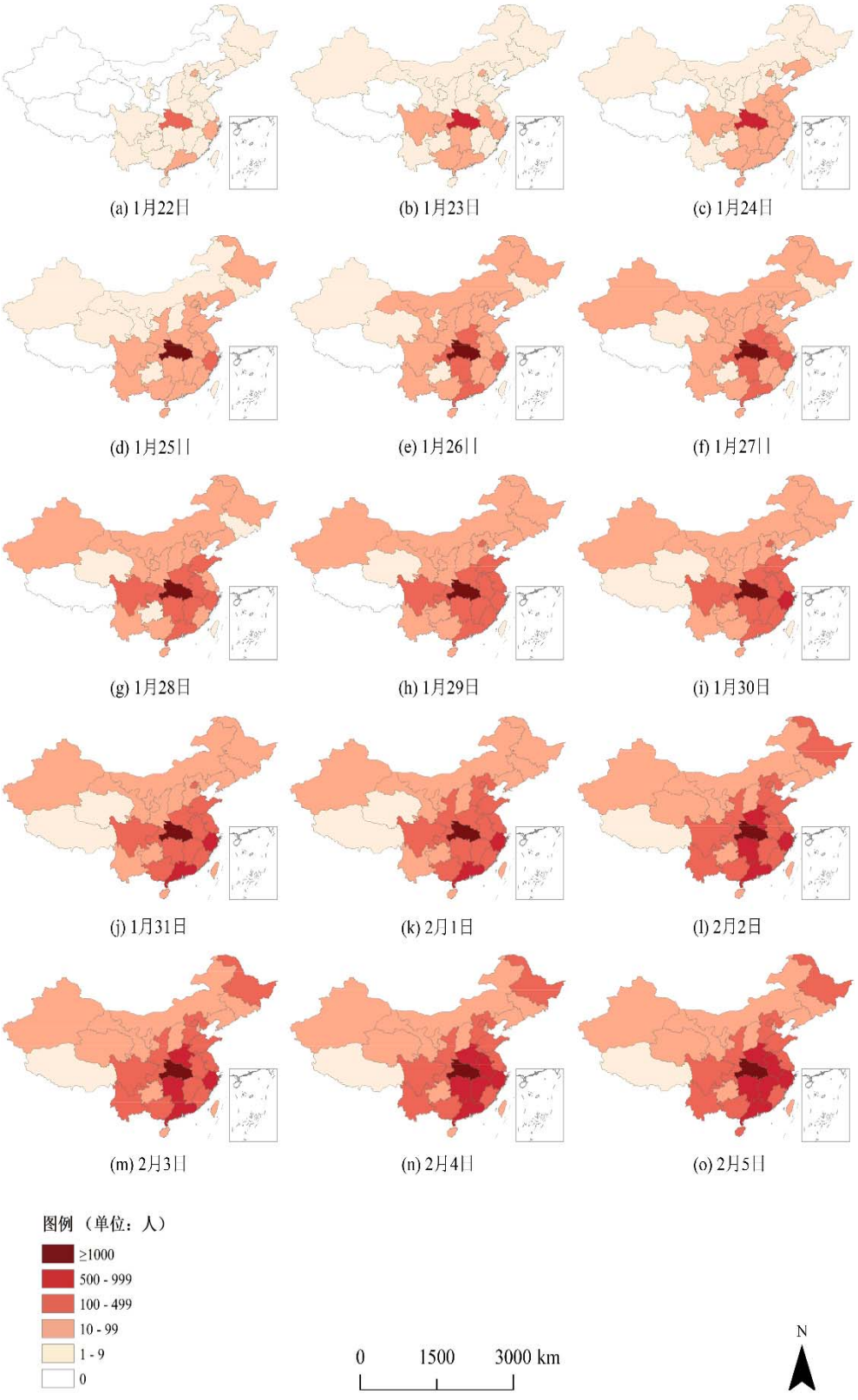

Figure 3 Temporal-spatial evolution map based on COVID-19 epidemic data (provincial scale)

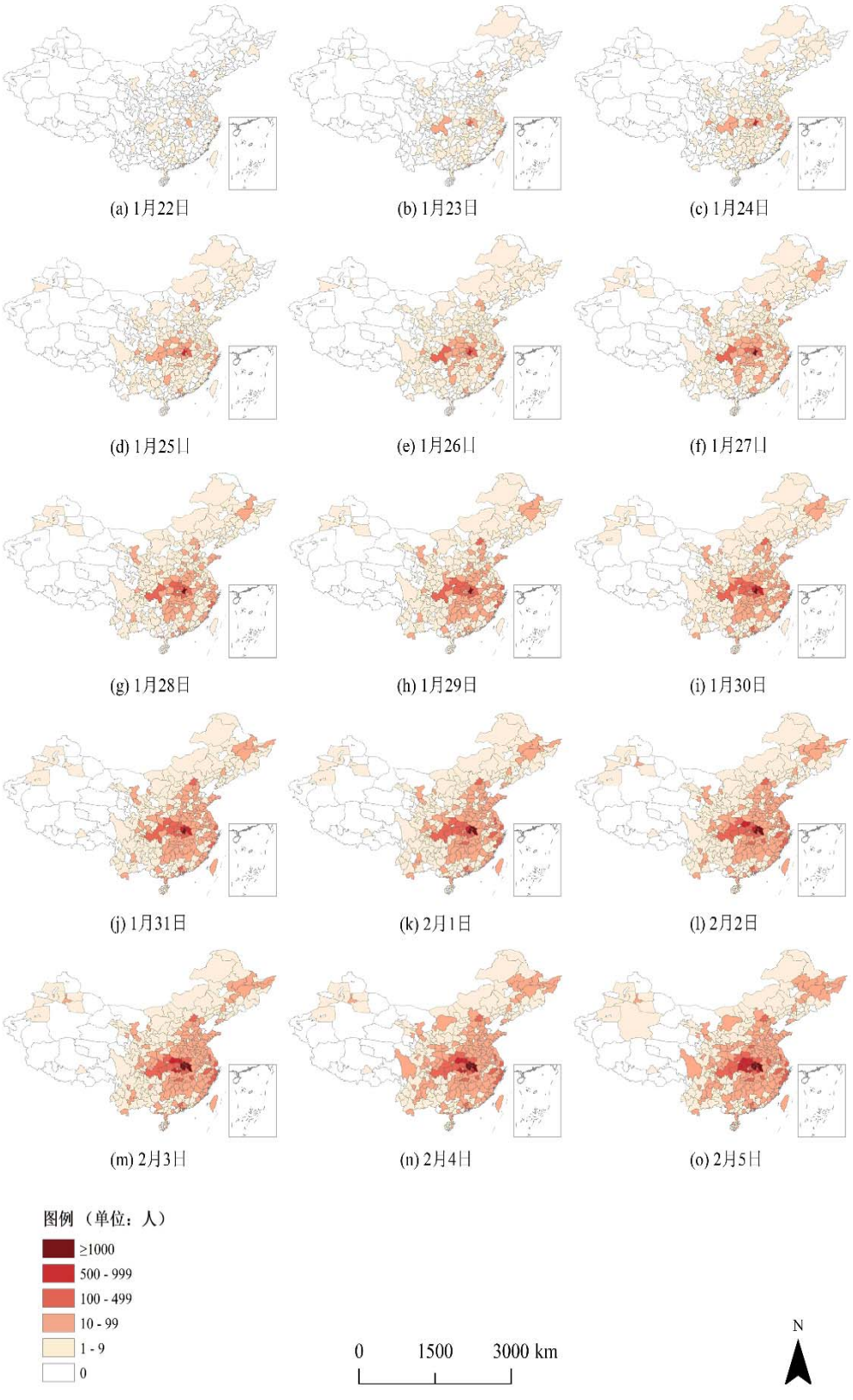

Figure 4 Temporal-spatial evolution map based on COVID-19 epidemic data (municipal scale)

5. Production Team

Junya Bao, Graduate student, School of Geography and Information Engineering, China University of Geosciences (Wuhan)

Hao Lin, Graduate student, School of Geography and Information Engineering, China University of Geosciences (Wuhan)

Qi Zhou, Associate Professor, School of Geography and Information Engineering, China University of Geosciences (Wuhan)

6. Contact Us

Junya Bao

Email address：estella970217@gmail.com

How to cite:

Bao, J, Lin, H., Zhou, Q., (2020). COVID-19 Epidemic Spatio-Temporal Dataset User Guide. https://github.com/Estelle0217/COVID-19-Epidemic-Dataset.git, 2020-02-18.