Flood Disaster Loss Dataset in China (2018)

Data Documentation

I. Dataset/atlas content features

i. Abstract

This dataset is Flood Disaster Loss Dataset in China (2018). Using web crawlers, We gathered professional reports related to earthquakes from the website of China National Commission for Disaster Reduction (NCDR-China). A series of extraction rules were constructed to extract disaster loss data. The following information was extracted: disaster time, event title, location, number of deaths, number of missing persons, affected population, direct economic losses, and crop disaster area. The data format is Excel. The spatial scope is China. The temporal range 2018.

ii. Elements (content fields)

There are totally 8 fields in the dataset, whose meaning is as follows:

“time”: when flood happended

“flood event”: flood describtion

“location” : flood loaction

“death”: number of deaths,

“mising”: number of missing persons,

“affected”: affected population,

“loss”: direct economic losses,

“crop”: crop disaster area

iii. Temporal cover

2018

iv. Spatial cover

The spatial range of this dataset is China.

II. Subject/industry scope of dataset/atlas

i. Subject scope

Geography .

ii. Industry scope

Disaster Risk Reduction

iii. Other classifications (optional)

(Other categories can be applied, but should reflect the dataset/atlas characteristics.)

III. Accuracy of dataset/atlas

i. Time frequency

ii. Spatial reference, accuracy, and granularity

IV. Dataset/atlas storage management

i. Data quantity

22.1KB

ii. Type format

Excel

iii. Update management

Irregular updating

V. Quality control of the dataset/atlas

i. Production mode

Data acquisition and preprocessing are carried out by means of network crawler, regular expression and natural language processing.

ii. Data sources (condition selection)

The dataset is collected from the website of China National Commission for Disaster Reduction (NCDR-China)

iii. Methods of the data acquisition and processing (condition selection)

Acquisition method: We gathered professional reports related to earthquakes from the website of China National Commission for Disaster Reduction (NCDR-China), which is one of the leading institutions that provides support to the government in addressing disaster-related issues by focusing on the entire cycle of disaster management. The “latest disaster” (“最新灾情” in Chinese) column of the website is selected as seed pages for the crawler. Using the Beautiful Soup Library in Python, we parsed the pages and collected 1,062 professional reports about natural disasters from January 1, 2015, to November 31, 2017. Of these, 219 were related to earthquakes.

Processing method: Rule-based and statistical learning-based methods are commonly used for information extraction. Due to the standardization of the language in professional report, we use a rule-based method for information extraction of flood information. Supported by the NLPIR-ICTCLAS Chinese lexical analysis system, a set of extraction rules relevant to flood is built, including temporal extraction rules, a location trigger dictionary, and an attribute trigger dictionary.

VI. Sharing and usage method of the dataset/atlas

i. Sharing methods and restrictions

Fully opened sharing

ii. Contact information of the sharing service (condition selection)

Contact Information for Service：

Name: Service group of Disaster Risk Reduction Knowledge Service System of IKCEST

Address：A11 Datun Road, Chaoyang District, Beijing

Zip Code：100101

E-mail: ikcest-drr@lreis.ac.cn

iii. Conditions and methods of usage

The dataset can be read by ArcGIS software and Microsoft office.

VII. Intellectual property rights of the dataset/atlas

i. Property rights (optional)

The property of the dataset belongs to the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences.

ii. Reference method of the dataset/atlas

Flood Disaster Loss Dataset in China (2018). Disaster Risk Reduction Knowledge Service of International Knowledge Centre for Engineering Sciences and Technology (IKCEST) under the Auspices of UNESCO, 2019.9.15.

iii. Usage contacts of the datasets/atlas

Name: Service group of Disaster Risk Reduction Knowledge Service System of IKCEST

Address: A11 Datun Road, Chaoyang District, Beijing .

Postcode: 100101

Telephone: 010-64889048-8006

Email: ikcest-drr@lreis.ac.cn

VIII. Others (optional)

In addition to the above, other information must also be explained.

Data documentation author information				
Data documentation author	HanXuehua	Update time		2019-09-25
Organization	Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences.			
Contact information				
Address	A11 Datun Road, Chaoyang District, Beijing .		PostcodeS	100101
Telephone	010-64889048-8006	E-mail	hanxh@lreis.ac.cn	