A METHOD FOR RECOGNIZING RAINFALL-SENSITIVE URBAN ROADS BASED ON TRAJECTORY DATA

Shaonan Zhu[1], Hongping Zhang[2*], You Jiang[3], Xin Yang[3]

[1]School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China - zhushaonan@njupt.edu.cn
[2]National Geomatics Center of China, Beijing, China-zhanghongping@ngcc.cn
[3]Key Laboratory of VGE of Ministry of Education, Nanjing Normal University, Nanjing, China

Commission IV, WG IV/3

KEY WORDS: Rainfall, Trajectory Data, Road Traffic, Mann-Kendall, Sensitivity Analysis, Urban Flooding.

ABSTRACT:
Rainfall has a substantial impact on urban traffic. Due to climate change and urbanization, rain and associated flooding will become an increasing hazard to urban traffic. Therefore, it is significant to detect road traffic anomalies and analyze their evolution. This study focuses on rainfall-sensitive roads, which have significant changes in speed and an apparent trend of change during multiple rainfalls. Firstly, the trajectory data are converted into road traffic time-series data, and the box plot is applied to determine whether there are significant outliers in roads speed. Secondly, the Mann-Kendall Trend Test is used to judge whether there is a noticeable trend change in the rainfall process. The proposed method is applied to Nanjing City in China. Based on the trajectory data from June to August 2019, the percentage of urban roads with significant changes in rainfall events is between 30% and 50%, with 217 roads that reveal significant changes every time. A total of 24 rainfall-sensitive roads are extracted by the trend test. This study will provide additional assistance for crucial road monitoring of urban rainstorm hazards.

1. INTRODUCTION

Rainfall has a significant influence on urban traffic (Li Q. et al., 2017). In particular, extreme rainfall typically causes urban flooding, leading to urban traffic disruptions, economic losses, and even casualties. In addition, due to climate change and urbanization, the frequency and intensity of extreme rainfall events are increasing (Huang, Cervone, and Zhang, 2017). It is foreseeable that excessive rainfall and urban flooding will become an increasing hazard to urban traffic. Therefore, it is significant to detect road traffic anomalies and analyze their evolution during rainfall events.

Related studies are focused on road traffic exposure under rainfall scenarios (Pyatkova et al., 2019). Rainfall and hydrodynamic models simulate different levels of rainfall and urban flooding(Li M. et al., 2018). The exposure of the overall road network is analyzed based on characteristics of urban road traffic such as speed and capacity. Yin et al.(2016) use high-resolution 2D inundation modeling to evaluate the risk of a pluvial flash flood on the road network in the city center of Shanghai, China. Zhu et al.(2018) predict the spatiotemporal distribution of flooding based on the Storm Water Management Model (SWMM) and TELAMAC-2D, use an agent-based model (ABM) to simulate driver behavior during a period of urban flooding, and find the effects with a 50-year return period were evident in the urban area of Lishui, China. Su et al. (2016) proposed an integrated simulation method for flood and traffic congestion under urban rainstorms based on the flood simulation model and questionnaire survey for human behavior. Such methods are based on rainfall, drainage network, and others, and can quantitatively analyze the spatial distribution of urban flooding. However, due to the complexity of the urban environment and drainage system, the accuracy of the flood model faces more significant challenges in practical applications.

With the emergence of trajectory data, the shortcomings of traditional traffic data collection are changed(Zheng, 2015). The rich trajectory data provides the basic data for in-depth exploration. It can help to identify road congestion (Belhadi et al., 2020) and analyze the pattern of operation of road traffic (Gong et al., 2020). Liu et al. (2013) propose a novel mobility-based approach clustering to identify different crowdedness spots, and verified by one-year GPS trajectory data. Zhang et al. (2020) develop a multi-task learning framework for Geodemographic inference by transit smart card data and provide new insights into understanding the relationship between human activity patterns and demographics. Chen (Chen, Jiang, and Sun, 2020) performed spatial and temporal correlation analysis on vehicle data to determine the congestion status of urban traffic and predicted the short-term traffic status using a mixed forest prediction method, which effectively alleviated the formation of traffic congestion. New approaches such as machine learning, deep learning, and spatial analysis in GIS, combined with the massive amount of trajectory data, produce various research results.

Several new studies have begun to apply trajectory data to road traffic under rainfall conditions. She et al.(2019) use partition statistics to construct a time series of GPS trajectory point density for each road, and extract the flooding roads by fusing GPS Trajectory data and Road Network. Ding et al.(2017) develop a weather-traffic index (WTI) system to identify regional weather-traffic sensitivity index throughout a city and analyze urban regions with a high impact of weather change on transport. Ni et al. (2021) collected 30 groups of scenarios with different levels of rainfall intensity and water depth to study the effect of a heavy rainstorm on urban traffic flow. With available data and technology, it is possible to support the discovery of roads with high exposure to rainstorm weather.

However, the urban system is dynamic, and rainstorm events also have spatial heterogeneity. Modeling and analyzing each rainstorm event is time-consuming work. City managers tend to focus more on finding "repeat offenders." So, this study proposes the concept of rainfall-sensitive roads, that is, the urban road with significant changes in speed and a clear trend of change during multiple rainfall events. We propose a convenient and efficient method to find roads with high exposure and high frequency in rainstorm weather.

The rest of the paper is organized as follows. The next section describes the study area and data source. The methods for abnormal road detection and rainfall-sensitive roads recognition are presented in Section 3. Section 4 discusses the results. The final section summarizes the essential findings and discusses future work.

2. STUDY AREA AND DATA SOURCE

The proposed method is applied to the center of Nanjing City in China. The experimental data include traffic data, rainfall data, and social sensing data.

Traffic data involves taxi trajectory point data and road network data. The trajectory data we obtained from the Nanjing traffic information center covers the period from June to August 2019. Road network data comes from OpenStreetMap (www.openstreetmap.org/). The trajectory dataset records the 24-hour operation of about 5000 taxis in Nanjing City, with a daily data volume of 40-50 million. It contains each vehicle on the road in detail, such as ID, longitude, latitude, driving speed, azimuth, and other information.

Rainfall data comes from rainfall monitoring stations. Six heavy rain (daily rainfall amounts over 25mm) events from June to August 2019 are selected.

Social sensing data is from China's social media platform Weibo(www.weibo.com/). We use key words query to obtain the relevant data on rainfall days in Nanjing, which is used to verify and analyze the rainfall-sensitive roads extracted in this paper.

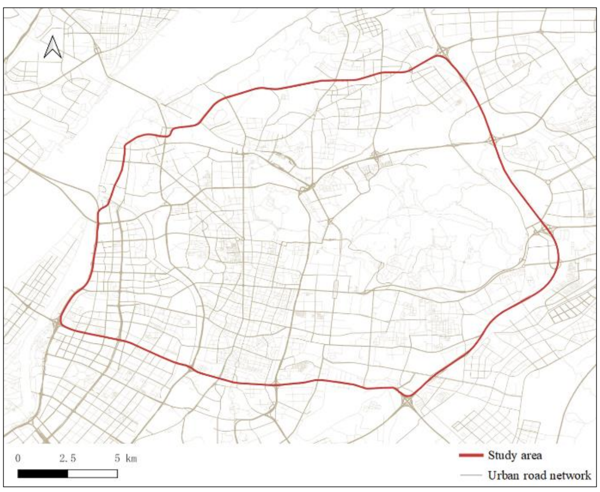

Figure 1. The study area in Nanjing City

Date	Cumulative (mm)	Duration (h)	Start time
2019/6/29	44.01	4	8:00
2019/7/1	63.90	3	19:00
2019/7/6	33.83	3	19:00
2019/7/12	34.57	3	17:00
2019/7/31	27.00	2	10:00
2019/8/10	52.77	2	7:00

Table 1. 6 rainfall events information

3. METHODOLOGY

3.1 Method Overview

The method mainly consists of two parts: speed variations significance test and rainfall-sensitive roads trend test. The former deals with multiple rainfall events, while the latter handles the rainfall process. Firstly, the trajectory data is converted into road traffic time-series data, and the box plot and the Speed Variation Rates (SVR) are applied to determine whether there are any significant outliers in roads speed during rainfall events. We use the Speed Variation Rates (SVR) to quantify the road speed discrepancy between rainfall and non-rainfall conditions. The average SVR under rainfall compares the average road speed under non-rainfall to recognize roads with obvious faster and slower speeds. Further, the roads with significant changes in all rainfall events are selected. Secondly, according to the results of the previous step, the Mann-Kendall Trend Test is used to judge whether there is a noticeable trend change during the rainfall process.

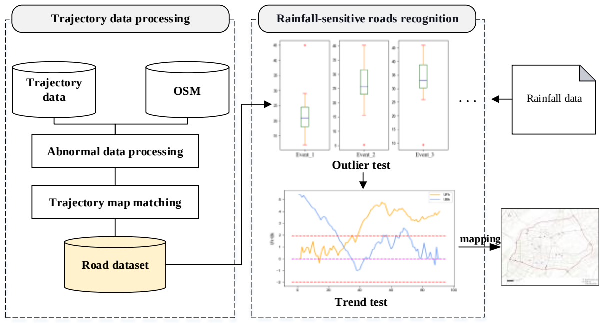

Figure 2. Flowchart of the methodology.

3.2 Speed variations significance test

The speed series is composed of the average speed of roads under non-rainfall weather and the road speed during rainfall. The box plot is constructed to determine whether obvious outliers exist in the speed of rainfall. Then, the Speed Variation Rates (SVR) (Li et al., 2017) is used to quantify the degree of the significant variations. The specific calculation formula is as follows:

$$x_i = \begin{cases} \dfrac{v_i^r - v_i^{up}}{v_i^{up}} (i = 1,2, \dots, \mathrm{n}), v_i^r > v_i^{up} \\ \dfrac{v_i^r - v_i^{low}}{v_i^{low}} (i = 1,2, \dots, \mathrm{n}), v_i^r < v_i^{low} \end{cases} \tag{1}$$

Where v_i^r is the speed of road *i* on non-rainfall days, v_i^{up} is the top edge of the box plot constructed by road *i*, v_i^{low} is the bottom edge of the box plot of road *i*. SVR values reflect whether the rainfall day will lead to distinct changes in road speed. Then, the rate of speed variations is classified to reflect the degree of road speed change. The average SVR is used as the calculation index, and the formula is:

$$\Delta v_{it} = \frac{\sum_{t=1}^{n} x_{it}}{n} \tag{2}$$

Where Δv_{it} is the average rate of change in speed of road *i* over n units of time. x_{it} is SVR of road *i* at time *t*. The type of road speed variation during rainfall can be classified into three categories by Δv_{it}: (1) $\Delta v_{it} > 0.1$ represents the road speed rise during rainfall; (2) $-0.1 \leq \Delta v_{it} \leq 0.1$ represents

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIII-B4-2022-217-2022

the road with no significant speed change during rainfall; (3) $\Delta v_{it} < -0.1$ means the road speed noticeable decline during rainfall.

3.3 Rainfall-sensitive roads trend test

Based on extracting the roads affected by rainfall, we further analyze whether there is a sudden trend of velocity change during the rainfall event. The Mann-Kendall test is applied to this study. The Mann-Kendall trend test (Mann, 1945; Kendall, 1975) is a non-parametric test method that is widely used to detect the clear trend of time series without considering the data distribution and can deal with outliers and missing values. Mann-Kendall test is applied to study whether there was a mutation trend in the road speed during rainfall. For the time series x containing n samples, the order column is constructed:

$$S_k = \sum_{i=1}^{k} r_i \; (k = 2,3,\dots,n) \tag{3}$$

Where:

$$r_i = \begin{cases} 1, x_i > x_j \\ 0, \; x_i \le x_j \end{cases} \; (j = 1,2,\dots,i) \tag{4}$$

$$E(S_k) = \frac{k(k-1)}{4} \; (2 \le k \le n) \tag{5}$$

$$Var(S_k) = \frac{k(k-1)(2k+5)}{72} (2 \le k \le n) \tag{6}$$

Assuming that time series data are random and independent, define statistics:

$$UF_k = \frac{[S_k - E(S_k)]}{\sqrt{Var(S_k)}} (k = 1,2,3,\dots,n) \tag{7}$$

Where UF_k is the standard normal distribution, which is the sequence data calculated by time series x in order of $(x_1, x_2, \cdots, x_n)$. At the same time, repeat the above calculation process according to the inverse order $(x_n, x_{n-1}, \cdots, x_1)$ of the time series x, while making $UF_k = UB_k (\mathrm{k} = \mathrm{n}, \mathrm{n}-1, \dots, 1)$ and $\mathrm{UB}_1 = 0$. If the values of both UF_k and UB_k statistics are exceeded 0 and the intersection point of their curves is within the range of the apparent horizontal line presenting the velocity-time series data showing an upward trend. Conversely, there is a continuing downward trend. However, if the intersection is outside the obvious local horizontal line, other methods need to determine whether there is a trend change.

4. EXPRIMENTAL AND RESULTS

4.1 Road speed variations under different rainfall events

The trajectory data are processed to obtain the time series data of each road speed. For example, in this stacked area graph, the upper edge line is the maximum speed of the Zhongyang road on non-rainfall days, the lower edge line is the minimum speed, the middle line is the average speed, and the remaining line indicates the road speed on the rainfall day (July 1, 2019), as shown in Figure 3.

It can be seen that the road speed on a rainy day is generally less than the average speed on a non-rainfall day. The speed at the moment of rainfall obviously deviates from the average value, the road speed is 0 at a period. After the rain, the road speed gets back up quickly and even exceeds the average speed.

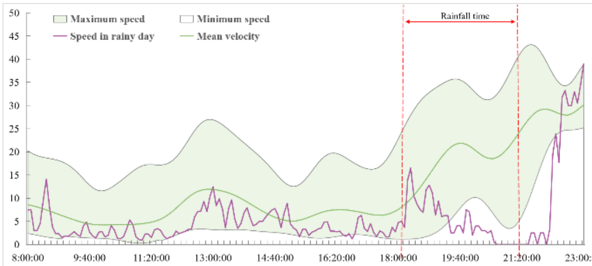

Figure 3. Comparison of road speed (From 8 a.m. to 11 p.m.)

Taking the rainfall time as the comparison range, the average speed of the non-rainfall days and the speed of the rainfall date are formed into a series. The box plot is used to determine whether there are abnormal data. In Figure 4, Zhongyang road speed anomalies are observed in all six rainfall events.

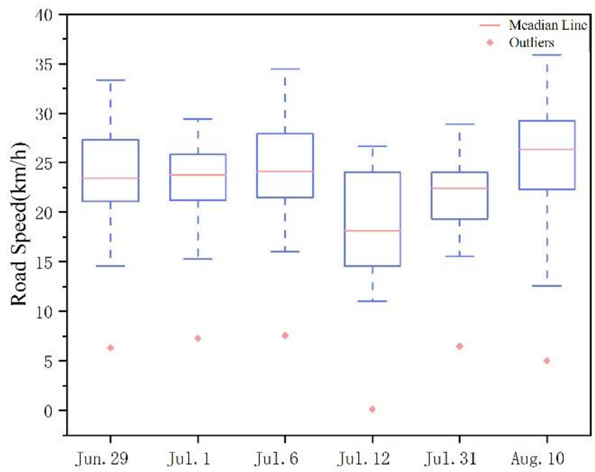

Figure 4. Box plot test of the Zhongyang road speed in rainfall events

Then, the box plot is used to test all roads with significant changes in the six rainfall events. The mean value of |t| and the percentage of significant roads among all roads are used to evaluate the impact of heavy rainfall on the entire urban road network. $|t| > 1.48$ implies that the roads change considerably in rainstorm weather when the significance level is $\alpha = 0.1$. As demonstrated in Table 2, at least 30 percent of roads are adversely impacted by rainstorms on any given rainy day. On July 1, the largest rainfall day, the proportion of significant roads reached 46.43%

Date	Average of $\lvert t \rvert$	Significant proportion
2019/6/29	1.67	45.32%
2019/7/1	1.40	46.43%
2019/7/6	1.62	32.64%
2019/7/12	1.72	35.65%
2019/7/31	1.68	32.12%

2019/8/10	1.71	31.65%

*Significance level $\alpha = 0.1,\ t_{\alpha}|n-1| = 1.48$.

Table 2. Results of box plot for all roads in rainfall events

SVR is used to further classify the significant roads. As Table 3 shown, all rainfall events cause a significant decrease in speed in about half of the significant roads. At the same time, about 10% of the road speed rise. It can be found that the intensity of rainfall and the start time of rainfall are two important influencing factors. In general, if rainfall is heavier, the impact will be more significant; if the rainfall intensity is comparable, rainfall in the morning and evening peak likewise leads to a significant slowdown.

Date	speed rise	non-significant	speed decline
2019/6/29	11.74%	35.60%	52.65%
2019/7/1	11.27%	27.60%	61.12%
2019/7/6	12.83%	33.22%	53.22%
2019/7/12	12.44%	37.59%	49.97%
2019/7/31	10.17%	37.86%	51.97%
2019/8/10	8.51%	45.28%	46.21%

Table 3. SVR calculation results

Figure 5 shows the spatial distribution of roads with significant speed decline in 6 rainfall events. A total of 217 roads are mapped. In terms of road types, primary, secondary, and tertiary dominate, totaling 90% of the total. Among them, secondary type roads account for the largest number of roads, reaching 78. From the spatial distribution, it is mainly concentrated in the city center (left side of the figure), while there are few roads in the northern part of the city. In terms of speed variation rates, most of them are less variable, while the large degree of variation presents a polycentric focus. Due to the small number of speed-rise roads in every rainfall event, no further study is conducted in this paper.

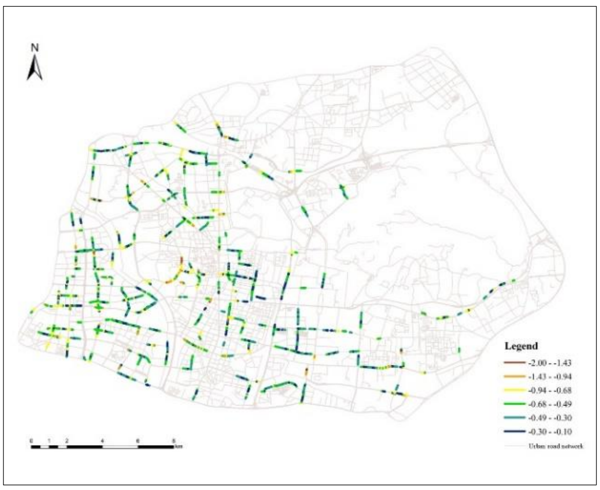

Figure 5. Spatial distribution of roads with significant speed decline in all rainfall events

4.2 Rainfall-sensitive roads recognition

The significant decline of road speed extracted in the previous step is used as input to the Mann-Kendall test, thus determining whether the road speed changed abruptly during the rainfall event. Figure 6 demonstrates the road speed variation(from 19:00 to 23:00 on July 1, 2019) for the road with ID=1857.

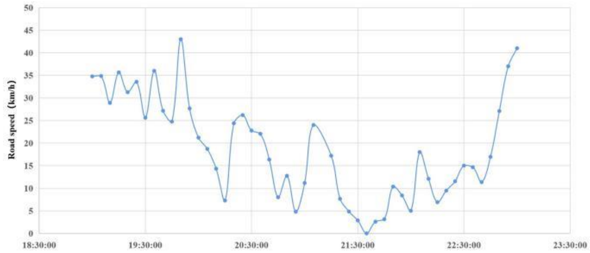

Figure 6. Road speed variations in one rainfall

Figure 7 is the Mann-Kendall trend test graph. It shows a decreasing trend after 19:30, and the road speed change abruptly. The intersection of the two lines coincides with the speed change in Figure 6, indicating that Mann-Kendall can detect the trend of the road change.

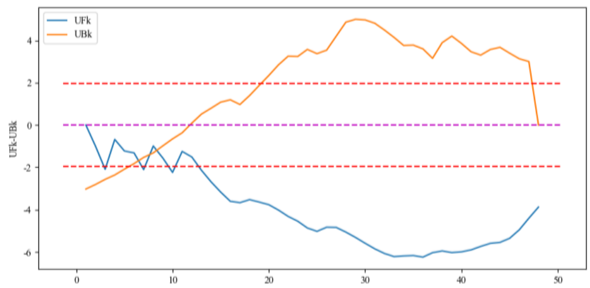

Figure 7. Trend test of road speed in one rainfall

This trend test method is used to calculate the road data in Figure 5, 24 roads are obtained, and the spatial distribution is shown in Figure 8. Comparison with Figure 5 indicates that rainfall-sensitive roads are concentrated in areas with high anomalies, and multiple centers are formed in the figure. On the other hand, the density of the road network is not directly related to the distribution of rainfall-sensitive roads.

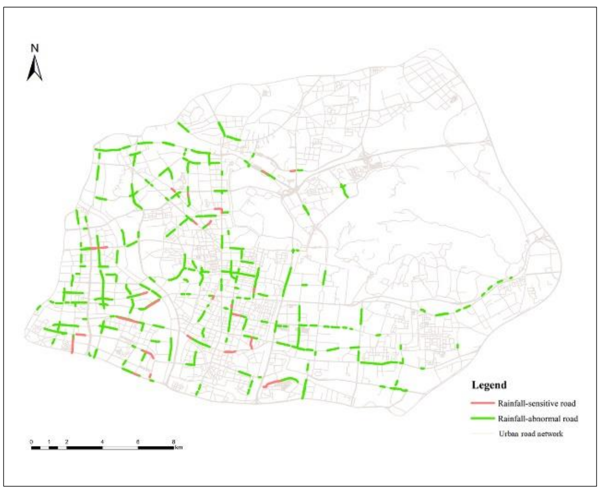

Figure 8. Spatial distribution of roads with speed variation trend

4.3 Results Analysis

Above, we mine roads in the urban transportation network that are sensitive to rainfall events from a quantitative perspective. The more interesting thing is what causes the sensitivity of the roads. So, social media (Weibo) data are used to conduct the analysis. The names of rainfall-sensitive roads are used as keywords for searching, and the dates of rainfall are used as time ranges for collecting relevant data. 246 relevant data are collected, and the data details are shown in Table 4.

Username	Content
Nanjing Traffic Radio	Flash flooding on Jiangdong Middle Road, north to south direction, near Yingtian Street; please pay attention.
****99	Zhongyang Road is flooding.
Jiangsu Traffic Radio Net	Provincial People's Hospital on Guangzhou Road is flooding, affecting the traffic flow.
Nanjing Morning Post	Mufu West Road has a large traffic flow in both directions.

Table 4. Example of Weibo data details

Further, the natural language technique is used to extract the place names, and we get the location by Geocoding. Both are from Baidu API. Thus, 40 roads related to waterlogging points are obtained in Figure 9. The number of rainfall-sensitive roads on or near the waterlogging point is 11 and is nearly half of the total. It is readily apparent that one of the main factors of rainfall-sensitive roads is urban flooding caused by heavy rainfall. On the contrary, the rainfall-sensitive roads test method in this paper has a specific ability to detect waterlogged roads in rainfall events.

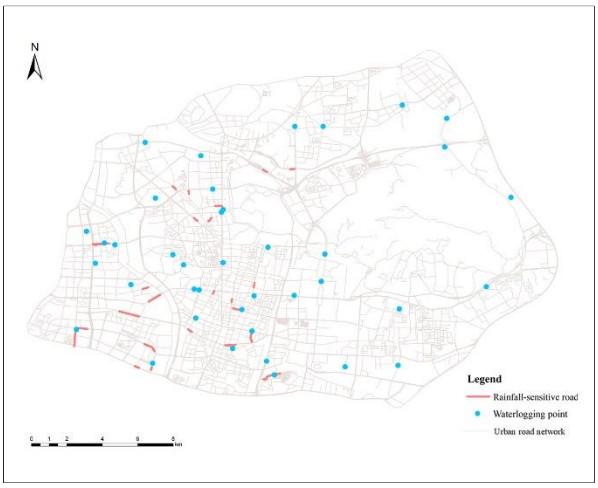

Figure 9. Spatial distribution of waterlogging points from Weibo

5. CONCLUSIONS

This study focuses on urban road speed variations during multiple rainfalls. The proposed method for recognizing rainfall-sensitive urban roads combines anomalous and speed trend tests based on trajectory data. Thus, trajectory data can provide a data basis for quantitatively mining the urban traffic operation patterns in rainfall events. However, due to the spatial heterogeneity of rainfall, the complexity of the urban system, and the mobility of the urban population, the factors considered in this paper are not comprehensive enough. In our future work, multiple data sources will be added to improve the accuracy of data processing, and the causes of rainfall-sensitive roads will be explored in depth.

REFERENCES

Belhadi, A., Djenouri, Y., Lin, J.C.-W., Cano, A., 2020. Trajectory Outlier Detection: Algorithms, Taxonomies, Evaluation, and Open Challenges. ACM Trans. Manage. Inf. Syst. 11, 16:1-16:29. doi.org/10.1145/3399631

Chen, Z., Jiang, Y., Sun, D., 2020. Discrimination and Prediction of Traffic Congestion States of Urban Road Network Based on Spatio-Temporal Correlation. *IEEE Access* 8, 3330–3342. doi.org/10.1109/ACCESS.2019.2959125

Ding, Y., Li, Y., Deng, K., Tan, H., Yuan, M., Ni, L.M., 2017. Detecting and Analyzing Urban Regions with High Impact of Weather Change on Transport. *IEEE Trans. Big Data* 3, 126–139. doi.org/10.1109/TBDATA.2016.2623320

Gong, S., Cartlidge, J., Bai, R., Yue, Y., Li, Q., Qiu, G., 2020. Extracting activity patterns from taxi trajectory data: a two-layer framework using spatio-temporal clustering, Bayesian probability and Monte Carlo simulation. *International Journal of Geographical Information Science* 34, 1210–1234. doi.org/10.1080/13658816.2019.1641715

Huang, Q., Cervone, G., Zhang, G., 2017. A cloud-enabled automatic disaster analysis system of multi-sourced data streams: An example synthesizing social media, remote sensing and Wikipedia data. *Computers, Environment and Urban Systems* 66, 23–37. doi.org/10.1016/j.compenvurbsys.2017.06.004

Kendall, M.G., 1975. Rank correlation methods. Griffin, Oxford, England.

Li M., Huang Q., Wang L., Yin J., Wang J., 2018. Modeling the traffic disruption caused by pluvial flash flood on intra-urban road network. *Transactions in GIS* 22, 311–322. doi.org/10.1111/tgis.12311

Li, Q., Hao, X., Wang, W., Wu, A., Xie, Z., 2017. EFFECTS OF THE RAINSTORM ON URBAN ROAD TRAFFIC SPEED — A CASE STUDY OF SHENZHEN, CHINA. Int. *Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W7, 71–75. doi.org/10.5194/isprs-archives-XLII-2-W7-71-2017

Liu, S., Liu, Y., Ni, L., Li, M., Fan, J., 2013. Detecting Crowdedness Spot in City Transportation. *IEEE Transactions on Vehicular Technology* 62, 1527–1539. doi.org/10.1109/TVT.2012.2231973

Mann, H.B., 1945. Nonparametric Tests Against Trend. *Econometrica* 13, 245–259. doi.org/10.2307/1907187

Ni, X., Huang, H., Chen, A., Liu, Y., Xing, H., 2021. Effect of Heavy Rainstorm and Rain-Induced Waterlogging on Traffic Flow on Urban Road Sections: *Integrated Experiment and Simulation Study. Journal of Transportation Engineering, Part A: Systems* 147, 04021057. doi.org/10.1061/JTEPBS.0000557

Pyatkova, K., Chen, A.S., Butler, D., Vojinović, Z., Djordjević, S., 2019. Assessing the knock-on effects of flooding on road transportation. *Journal of Environmental Management* 244, 48–60. doi.org/10.1016/j.jenvman.2019.05.013

She, S., Zhong, H., Fang, Z., Zheng, M., Zhou, Y., 2019. Extracting Flooded Roads by Fusing GPS Trajectories and Road Network. *ISPRS International Journal of Geo-Information* 8, 407. doi.org/10.3390/ijgi8090407

Su, B., Huang, H., Li, Y., 2016. Integrated simulation method for waterlogging and traffic congestion under urban rainstorms. *Nat Hazards* 81, 23–40. doi.org/10.1007/s11069-015-2064-4

Wu, C.-H., Ho, J.-M., Lee, D.T., 2004. Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems* 5, 276–281. doi.org/10.1109/TITS.2004.837813

Yin, J., Yu, D., Yin, Z., Liu, M., He, Q., 2016. Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of Shanghai, China. *Journal of Hydrology* 537, 138–145. doi.org/10.1016/j.jhydrol.2016.03.037

Zhang, Y., Sari Aslam, N., Lai, J., Cheng, T., 2020. You are how you travel: A multi-task learning framework for Geodemographic inference using transit smart card data. *Computers, Environment and Urban Systems* 83, 101517. doi.org/10.1016/j.compenvurbsys.2020.101517

Zheng, Y., 2015. Trajectory Data Mining: An Overview. *ACM Trans. Intell. Syst. Technol.* 6, 29:1-29:41. doi.org/10.1145/2743025

Zhu, J., Dai, Q., Deng, Y., Zhang, A., Zhang, Y., Zhang, S., 2018. Indirect Damage of Urban Flooding: Investigation of Flood-Induced Traffic Congestion Using Dynamic Modeling. *Water* 10, 622. doi.org/10.3390/w10050622