Journal Pre-proof

Investigating the Impact of the COVID-19 Pandemic on Crime Incidents Number in Different Cities

Miaomiao Hou , Zhaolong Zeng , Xiaofeng Hu , Jinming Hu

PII: S2666-4496(22)00013-5
DOI: https://doi.org/10.1016/j.jnlssr.2021.10.008
Reference: JNLSSR 66

To appear in: *Journal of Safety Science and Resilience*

Received date: 8 July 2021
Revised date: 18 October 2021
Accepted date: 19 October 2021

Please cite this article as: Miaomiao Hou , Zhaolong Zeng , Xiaofeng Hu , Jinming Hu , Investigating the Impact of the COVID-19 Pandemic on Crime Incidents Number in Different Cities, *Journal of Safety Science and Resilience* (2022), doi: https://doi.org/10.1016/j.jnlssr.2021.10.008

<PE-AT>Investigating the Impact of the COVID-19 Pandemic on Crime Incidents Number in Different Cities

Miaomiao Hou[a], Zhaolong Zeng[a, *], Xiaofeng Hu[a], Jinming Hu[a]

[a] School of Information Technology and Cybersecurity, People's Public Security University of China, Beijing 102628, China

*Corresponding author.

E-mail address:

1069735422@qq.com (Miaomiao Hou).

zengzhaolong@ppsuc.edu.cn (Zhaolong Zeng).

huxiaofeng@ppsuc.edu.cn (Xiaofeng Hu).

377024160@qq.com (Jinming Hu).

Abstract

The COVID-19 pandemic is strongly affecting many aspects of human life and society around the world. To investigate whether this pandemic also influences crime, the differences of crime incidents numbers before and during the pandemic in four large cities (namely Washington DC, Chicago, New York City and Los Angeles) are investigated. Moreover, the Granger causal relationships between crime incidents numbers and new cases of COVID-19 are also examined. Based on that, new cases of COVID-19 with significant Granger causal correlations are used to improve the crime prediction performance. The results show that crime is generally impacted by the COVID-19 pandemic, but it varies in different cities and with different crime types. Most types of crimes have seen fewer incidents numbers during the pandemic than before. Several Granger causal correlations are found between the COVID-19 cases and crime incidents in these cities. More specifically, crime incidents numbers of theft in DC, Chicago and New York City, fraud in DC and Los Angeles, assault in Chicago and New York City, and robbery in Los Angeles and New York City, are significantly Granger caused by the new case of COVID-19. These results may be partially explained by the Routine Activity theory and Opportunity theory that people may prefer to stay at home to avoid being infected with COVID-19 during the pandemic, giving fewer chances for crimes. In addition, involving new cases of COVID-19 as a variable can slightly improve the performance of crime prediction in terms of some specific types of crime. This study is expected to obtain deeper insights to the relationships between the pandemic and crime in different cities, and to provide new attempts for crime prediction during the pandemic.

Keywords: COVID-19 pandemic, Crime incidents numbers, Crime prediction, Granger causality, Long Short-Term Memory Network

<PE-FRONTEND>

1 Introduction

The COVID-19 pandemic is one of the most serious global public health events in recent years. The onset and spread of COVID-19 have affected nearly every continent. People’s daily lives and the whole society have been drastically influenced around the world[1-3]. For example, in many cities, traffic is completely restricted[4, 5]; non-essential businesses have closed for a very long time; travel became more and more difficult; and social gatherings are limited[6, 7]. Moreover, the COVID-19 pandemic is a huge challenge to education activities[8, 9], many courses are moved online. At the same time, unemployment among many groups of workers increased sharply[10, 11]. What's more serious is that the pandemic has led to a dramatic loss of human life, economic losses and social disruption worldwide, presenting an unprecedented challenge to public health, food systems, and public safety[12]. This also raises attention to other questions related to our lives and security. Does the COVID-19 pandemic have an impact on crime? If so, is this impact strong or weak? If the COVID-19 pandemic has a strong impact on crime, will the pandemic be a factor for analyzing and predicting crimes? These questions motivated this study.

During the COVID-19 pandemic, a few studies have investigated the impact of the pandemic on crime in different regions. For instance, Shayegh and Malpede[13] identified an overall drop by about 40% across crime types in San Francisco and Oakland from March 16, 2020 to March 28, 2020. Campedelli, Aziani and Favarin[14] conducted Bayesian structural time-series and focused on nine crime categories, identifying that overall crime has significantly decreased in Los Angeles, as well as robbery, shoplifting, theft, and battery. Felson, Jiang and Xu[15] examined burglary in Detroit during three periods which are related to government suggestions, their findings indicated an overall 32% decline in burglary, with the most substantial change in the third period. De la Miyar, Hoehn-Velasco and Silverio-Murillo[16] used an event study for the intertemporal variation across the 16 districts' eight common crimes in Mexico City for 2019 and 2020, and indicated a decline in conventional crime during the COVID- 19 pandemic, while organized crime remains steady. Ashby[17] found that burglary only declined in Austin, Los Angeles, Memphis, and Scan Francisco, serious assaults declined in Austin, Los Angeles, and Louisville, but not other cities.

In these previous studies, investigation of impact on fraud was not reported. And what's the difference of the impacts on Chinese cities was seldom studied. Furthermore, in terms of the time series of the daily crime incidents and COVID-19 cases, are there significant correlations in different cities? This is still an open question.

As is known to all, crimes are affected by many factors, such as economic variables[18-21], spatial and temporal autocorrelation factors[22-31], environmental conditions[32-37], and current politics[38, 39]. These variables are often used to predict crimes[22, 26, 40], providing support for crime risk prevention and control. Thus, another noteworthy question is that whether COVID-19 pandemic could be considered as a new factor to predict crime? Previous studies gave few ideas about that.

In this paper, we first investigate the differences of four common crime incidents numbers (theft, fraud, assault, robbery and burglary) in four large cities (namely Washington DC, Chicago, New York City and Los Angeles) before and during the pandemic. Then the Granger causal relationships between crime incidents numbers and new cases of COVID-19 are studied. Finally, based on the results of Granger causality test between crime and COVID-19 pandemic, new cases of COVID-19 with significant Granger causal correlations are conducted to crime prediction to examine whether the new cases can improve the prediction performance of the daily crime incidents numbers.

The paper is organized as follows: In “Introduction”, the research background is introduced. “Materials and Methods” describes the data sets used in this study including the new cases of COVID-19 and crime incidents number in different cities, and focuses on the theory and steps of the Granger causality test and crime prediction. The results are presented and discussed in “Results and discussion”. Finally, “Conclusion” draws a conclusion and points to the future research.

2 Data and Methods

2.1 Data description

Four large cities, namely Washington DC, Chicago, New York City and Los Angeles are selected as the research areas. These cities are typical large cities in the United States with adequate crime and COVID-19 data, and they have similar Economic, cultural and social background. Thus, it is reasonable to compare the impacts of COVID-19 cases (as well as people's activities influenced by them) on crime patterns among the above four cities. As for crime types, theft, fraud, assault, robbery and burglary are all the most common kinds of crimes through the world, specifically, theft, fraud and burglary are property crimes, while robbery and assault are violent crimes. The daily crime records of the US cities are taken from the Open Data DC in the Office of the Chief Technology Officer (https://opendata.dc.gov), the Chicago open data portal (https://data.cityofchicago.org), NYC Open Data (https://opendata.cityofnewyork.us) and Los Angeles Open Data (https://data.lacity.org), that contain the pandemic information collected by government organizations, and free download service. The new COVID-19 cases of the US cities are collected from the Bing COVID-19-Data GitHub repo (https://www.bing.com/covid).

From the datasets, we found that turning point existed, which can divide periods into those before and during the COVID-19 pandemic. Fig. 1 shows the daily new COVID-19 cases in the above cities in which the turning points of the pandemic are marked. For these cities, the turning points are assumed at the days which are related to government prevention and control recommendations of the COVID-19 pandemic. More specifically, the turning point of Washington DC is April 1, 2020[41], Chicago's turning point is March 21, 2020[42], New York City's turning point is March 20, 2020[43] and Los Angeles's turning point is March 15, 2020[44]. As the daily crime incidents numbers in New York City were updated untimely, the research period of the New York City during the pandemic only lasted until September 30, 2020. And the research periods lasted until November 30, 2020, in other cities. As shown in Fig. 1, the number of new COVID-19 cases in Washington DC is low and with less fluctuation. While in Chicago, the first COVID-19 pandemic wave is in April and May, and the second wave is in November. The number of new cases in New York City increases firstly, and then slowly decreases to a stable level. And the statistic is quite different in Los Angeles. The number of new cases increases to the first plateau in July, then slowly decreases, and finally breaks out rapidly in November.

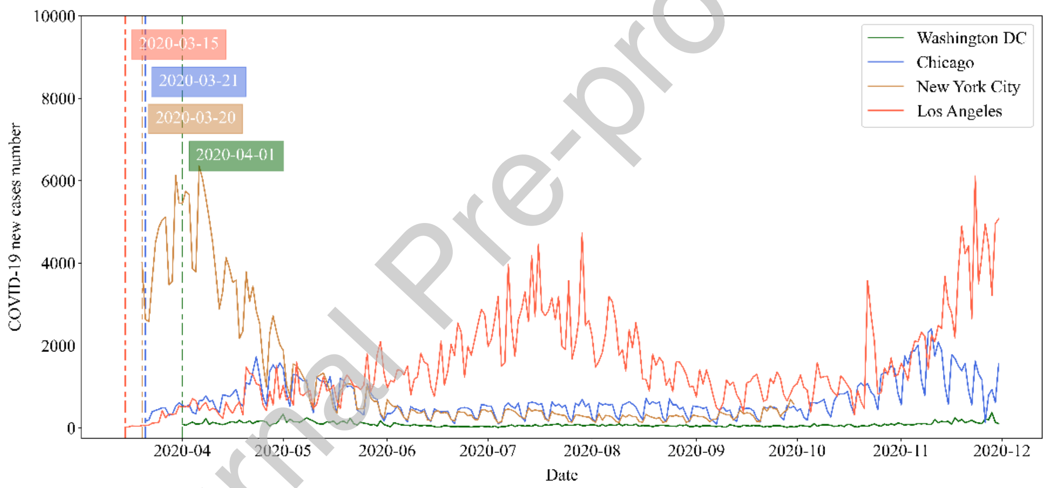

Figure 1. The daily new cases of COVID-19 in different cities. Green line, blue line, brown line and red line represent new confirmed cases of COVID-19 in Chicago, New York City and Los Angeles respectively. Green dotted line, blue dotted line, brown dotted line and red dotted line represent turning points of COVID-19 pandemic in DC, Chicago, New York City and Los Angeles respectively.

Daily crime incidences numbers in different cities are shown in Fig. 2. Daily crime incidents in DC are fewer than those in the other cities. It is shown that almost all types of crimes witnessed significant decreasing trend from 2020 to 2021. Also, seasonality and relatively steady daily variations of theft, assault, robbery and burglary can be observed in all the four cities. However for fraud, the daily incidents numbers fluctuate greatly and its seasonal cycle is not very clear as shown in the second panel. The descriptive statistics of crime incidents numbers before and during the pandemic in different cities are shown in Table 1. In these cities, theft and fraud incidents numbers are large and spread out over a wider range than the other types of crimes incidents numbers in most of the cases. For example, the number of theft incidents (M= 44.828, SD= 7.894) is larger than that of fraud incidents (M= 28.955, SD= 8.532), assault incidents (M= 4.631, SD= 2.258), robbery incidents (M= 6.643, SD= 2.906) and burglary incidents (M= 3.398, SD= 1.963) in DC before the pandemic. Here, M is the mean value of the crime incidents number and SD indicates the standard deviation. In addition, some crime incidents numbers during the pandemic are much fewer than those before the pandemic in these cities, such as theft.

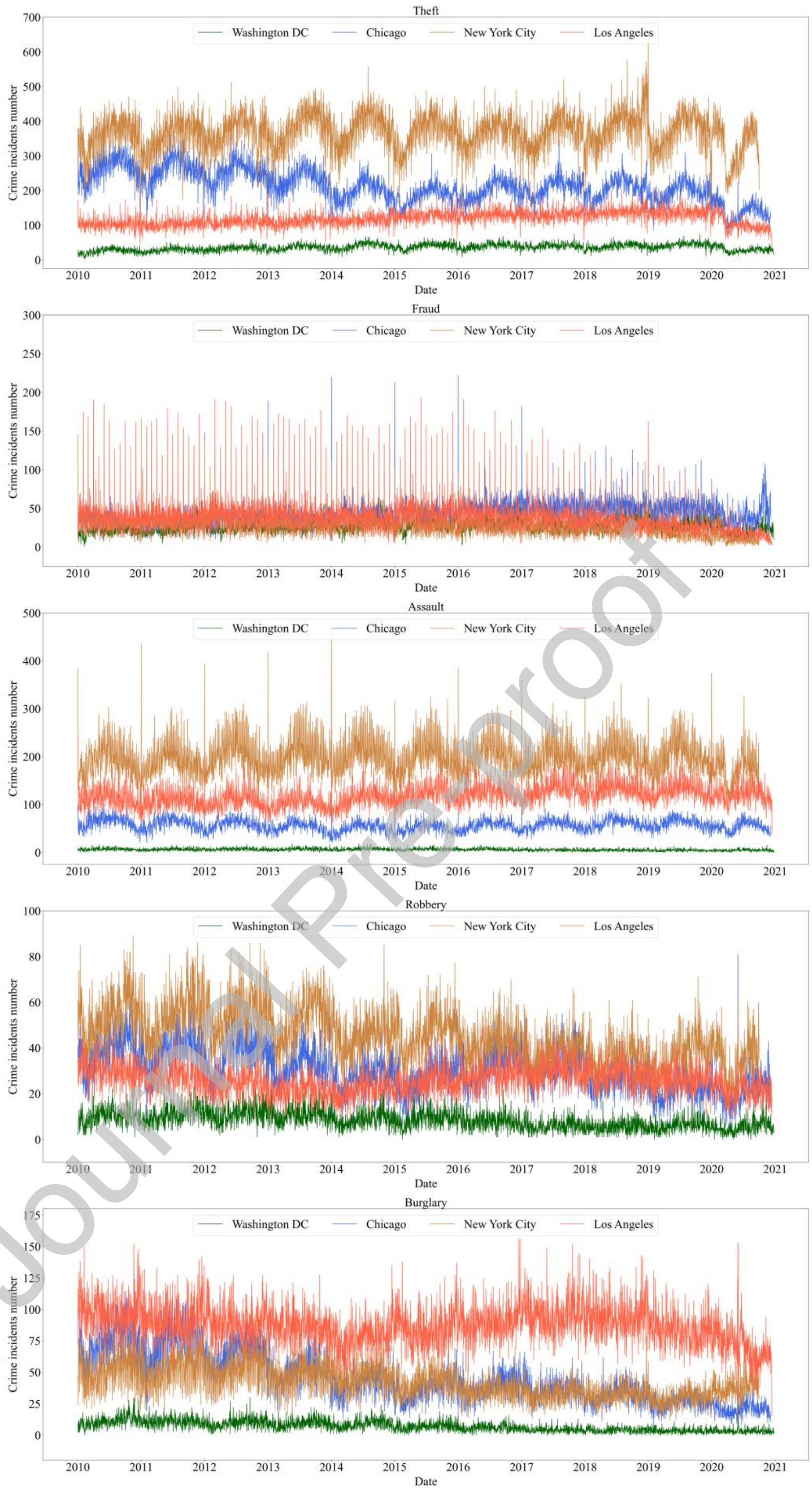

Figure 2. Daily crime incidences numbers in different cities. Green, blue, brown and red line represent the daily crime incidents numbers in DC, Chicago, New York City and Los Angeles, respectively.

Table 1. The descriptive statistics of crime incidents numbers before and during the pandemic in different cities.

City	The research period	Data category	Mean	Standard deviation
Washington DC	Before the pandemic (From April 1, 2019 to November 30, 2019)	Theft	44.828	7.894
		Fraud	28.955	8.532
		Assault	4.631	2.258
		Robbery	6.643	2.906
		Burglary	3.398	1.963
	During the pandemic (From April 1, 2020 to November 30, 2020)	Theft	27.020	6.866
		Fraud	20.320	6.982
		Assault	4.775	2.385
		Robbery	5.857	3.184
		Burglary	3.381	2.626
Chicago	Before the pandemic (From March 21, 2019 to November 30, 2019)	Theft	202.388	27.107
		Fraud	51.380	13.658
		Assault	63.886	10.560
		Robbery	22.592	5.888
		Burglary	27.125	5.920
	During the pandemic (From March 21, 2020 to November 30, 2020)	Theft	132.792	24.017
		Fraud	40.576	16.910
		Assault	53.925	11.785
		Robbery	21.933	7.757
		Burglary	21.164	8.862
New York City	Before the pandemic (From March 20, 20219 to September 30, 2019)	Theft	386.210	45.522
		Fraud	17.005	6.163
		Assault	215.523	33.840
		Robbery	37.287	8.316
		Burglary	29.400	6.953
	During the pandemic (From March 20, 2020 to September 30, 2020)	Theft	309.907	60.942
		Fraud	8.538	4.213
		Assault	170.256	36.746
		Robbery	32.656	10.195
		Burglary	39.579	9.390
Los Angeles	Before the pandemic (From March 15, 2019 to November 30, 2019)	Theft	136.602	17.411
		Fraud	28.068	13.287
		Assault	133.602	21.040
		Robbery	26.368	5.829
		Burglary	82.330	12.541
	During the pandemic (From March 15, 2020 to November 30, 2020)	Theft	94.85	14.218
		Fraud	17.70	7.370
		Assault	120.64	20.594

Robbery	21.23	5.206
Burglary	68.52	15.351

2.2 The Granger causality test between COVID-19 pandemic and crime

To know whether the COVID-19 pandemic influence crimes, Granger causality test is applied[45]. First, time series stationarity are examined with the Augmented Dickey-Fuller (ADF) unit root test to avoid spurious regression[46]. If the calculated ADF statistic is lower than 1% and the *p* values of the significance level is lower than 0.05, the null hypothesis that assumes the presence of unit root is rejected, inferring that the time series is stationary. In contrast, if the null hypothesis is not rejected, the time series should be non-stationary. In this study, the first-order difference is applied to non-stationary time series to make all the time series stationary. Then, the Granger causality test is performed, and the optimal lags of the Granger causality test are obtained by the vector autoregressive models through the minimum Akaike information criterion (AIC) value[47].

The essence of the Granger causality test is to test whether the lagged values of a time series can be introduced into the equation of other time series. If a time series is influenced by the lagged values of other time series, both series have Granger causality. For crime time series Y_t and new COVID-19 cases time series X_t, the regression equation is represented as follows:

$$Y_t = \sum_{i=1}^{k} \alpha_i Y_{t-i} + \sum_{i=1}^{k} \beta_i X_{t-i} + e_t \qquad [40]$$

where k represents the number of lags included in the regression, α_i and β_i represents the weights of Y_{t-i} and X_{t-i}, and e_t represents random white noise. The null hypothesis of time series Y_t and new COVID-19 cases time series X_t is COVID-19 pandemic does not Granger cause crime, namely, H_0: β_i =1 (i=1, 2, …, k), the test statistic for the null hypothesis is computed as follows:

$$F = \frac{(RSS_R - RSS_U)/k}{RSS_U/(T-2k-1)} \sim F(k, T-2k-1) \tag{2}$$

where T is the sample size, RSS_R represents the residual sum of squares of Eq. [40] when $\beta_i = 1$ (i=1, 2, …, k), and RSS_U represents the residual sum of squares of equation [40] when $\beta_i \neq 1$ (i=1, 2, …, k). The test statistic follows an F distribution with k and $T-2k-1$ degrees of freedom. If the p value of F-Statistic is lower than 0.05, we can reject the null hypothesis which means that new confirmed cases of COVID-19 Granger cause the crime incidents numbers. In contrast, we accept the null hypothesis, and there is no Granger causal relationship between crime incidents numbers and new confirmed cases of COVID-19.

2.3 Crime prediction based on LSTM

Based on the results of the Granger causality test, several Granger causal relationships are confirmed. For each pair of COVID-19 cases and crime with significant Granger causal relationship, research on crime prediction is implemented. Here, the new cases of COVID-19 are treated as a new feature for the crime prediction model. In this study, Long Short-Term Memory (LSTM) model is used to predict crimes and to examine whether the new cases of COVID-19 can improve the performance of the prediction for the daily crime incidents numbers.

LSTM is an improved multilayer perceptron network based on Recurrent Neural Network (RNN) which is widely used for time series prediction[48]. LSTM adds a memory unit to each hidden layer neural unit to realize controllable memory information in time series. When the time series data is transferred between the units of the hidden layer, it will pass through the input gate, forget gate, output gate and other interactive controllable gates to control the memory of previous data and current data, and the degree of forgetting, so that the neural network has a long-term memory function. In this way, LSTM effectively overcomes shortcomings such as the traditional

RNN gradient disappearance, defects in effectively retaining long-term memory information[49]. In this study, new cases of COVID-19 with significant Granger causal correlations are applied in the LSTM models to improve the crime prediction performance. The time series are divided as shown in Table 2. The test set is the last two weeks of the time series (14 days). The rest of the time series is the train set.

Table 2. The division of the time series.

City	Train Sets	Test Sets
Washington DC	From 1/1/2010 to 11/16/2020	From 11/17/2020 to 11/30/2020
Chicago	From 1/12010 to 11/162020	From 11/17/2020 to 11/30/2020
New York City	From 1/1/2010 to 9/16/2020	From 9/17/2020 to 9/30/2020
Los Angeles	From 1/1/2010 to 11/16/2020	From 11/17/2020 to 11/30/2020

The statistical law of stationary time series data changes little over time, and can usually be used for time series prediction. Therefore, stationarity of the time series is examined by ADF test and can apply the first-order difference to make all the time series stationary firstly. Some features including "month", "weekend", "holiday", "weekday_avg", "weekend_avg" and "month_avg" are extracted in this study (see Table 3). Next, the number of lagging observations is set to one. In other words, the crime incidents numbers at the previous moment are used to predict the crime incidents numbers at the current moment. Finally, all the time series are normalized for LSTM model training.

Table 3. Data Features.

Feature name	Feature value
month	Current month
weekend	0: weekday, 1: weekend
holiday	0: non-holiday, 1: holiday
weekday_avg	Average number of crime incidents per weekday
weekend_avg	Average number of crimes incidents per weekend
month_avg	Average number of crimes incidents per month

The LSTM models constructed in this paper is mainly composed of two layers: LSTM and Dense. The train sets that do not contain the features of daily new cases are inputted to the LSTM model. The model uses RMSE as the loss function and uses the Adaptive Moment Estimation to optimize it. The prediction results without conducting the new COVID-19 cases are obtained by input the test set into the LSTM models. Then, the train sets that contain the features of the COVID-19 pandemic are inputted to the LSTM model with the same parameters. The prediction results with conducting new cases are obtained by inputting the test set into the LSTM models. To examine whether the new cases of COVID-19 can improve the prediction performance of the daily crime incidents numbers, two indices are used to quantitatively evaluate and compare the prediction results with and without conducting the new COVID-19 cases: root mean square error (RMSE)[31], and percentage root mean square error (PRMSE)[50], which are defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(O_t - P_t)^2} \times 100\% \tag{3}$$

$$PRMSE = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(O_t - P_t)^2}}{\frac{1}{n}\sum_{t=1}^{n}O_t} \times 100\% \tag{4}$$

where O_t represents the observed value and P_t represents the predicted value. n is the total number of predicted days, and the value of n in this study is 14.

3 Results and discussion

3.1 The difference of crime incidents numbers before and during the COVID-19 pandemic

To examine the impact of COVID-19 on crime incidents numbers in large cities, the differences of crime incidents numbers before and during the COVID-19 pandemic are investigated. Fig. 3 shows the distributions of daily theft incidents numbers before and during the pandemic in four large cities of the U.S (Washington DC, Chicago, New York City and Los Angeles). As shown in Fig. 3, the theft incidents numbers during the pandemic are much fewer than those before the pandemic in all the four cities. The stories of fraud, assault and robbery are quite like that of theft which means that all the four selected cities in the US witness significant decreases of crime incidents number (as shown in Fig. A1, Fig. A2 and Fig. A3, respectively).

The Mobility Trends Reports of Apple (see https://covid19.apple.com/mobility) provides the relative volume of route requests for each country/region, subregion, or city compared to the baseline volume on January 13, 2020. The data of Apple Mobility Trends Reports is based on the direction requested by the users in Apple Maps, which are classified into three categories: walking, driving, and public transit. The average of three categories' relative volume is selected to represent mobile trends in different cities. Fig. 4 shows the mobile trends before and during the COVID-19 pandemic in different cities of the U.S. As shown in Fig. 4, the mobile trends of these four cities decreased significantly in March. Specifically, the mobile trend of New York City on March 29, 2020 decreased by 78.97% compared to the baseline. This indicates that people's activities may have greatly reduced since the COVID-19 pandemic. Because the local authorities of these cities imposed a range of strategies during the COVID-19 pandemic, such as stay-at-home orders, travel bans, closures of schools and so on[51]. These strategies aimed at limiting interaction to avoid being infected by COVID-19 during the pandemic. According to Routine Activity (RA) theory, less person-to-person contact means less opportunity for crimes, which may explain why the incidents numbers of crimes decreased during the pandemic.

An exception is burglary in New York City, since obvious increase of incidents number is witnessed in Fig. A4, while in DC, Chicago and Los Angeles the trend is decreasing. As reported by the New York Post (see https://nypost.com/2020/11/14/new-stats-reveal-massive-nyc-exodus-amid-coronavirus-crime/), more than 300,000 New Yorkers moved away from the city during the pandemic, so large numbers of unoccupied houses and departments may provide many chances for burglaries.

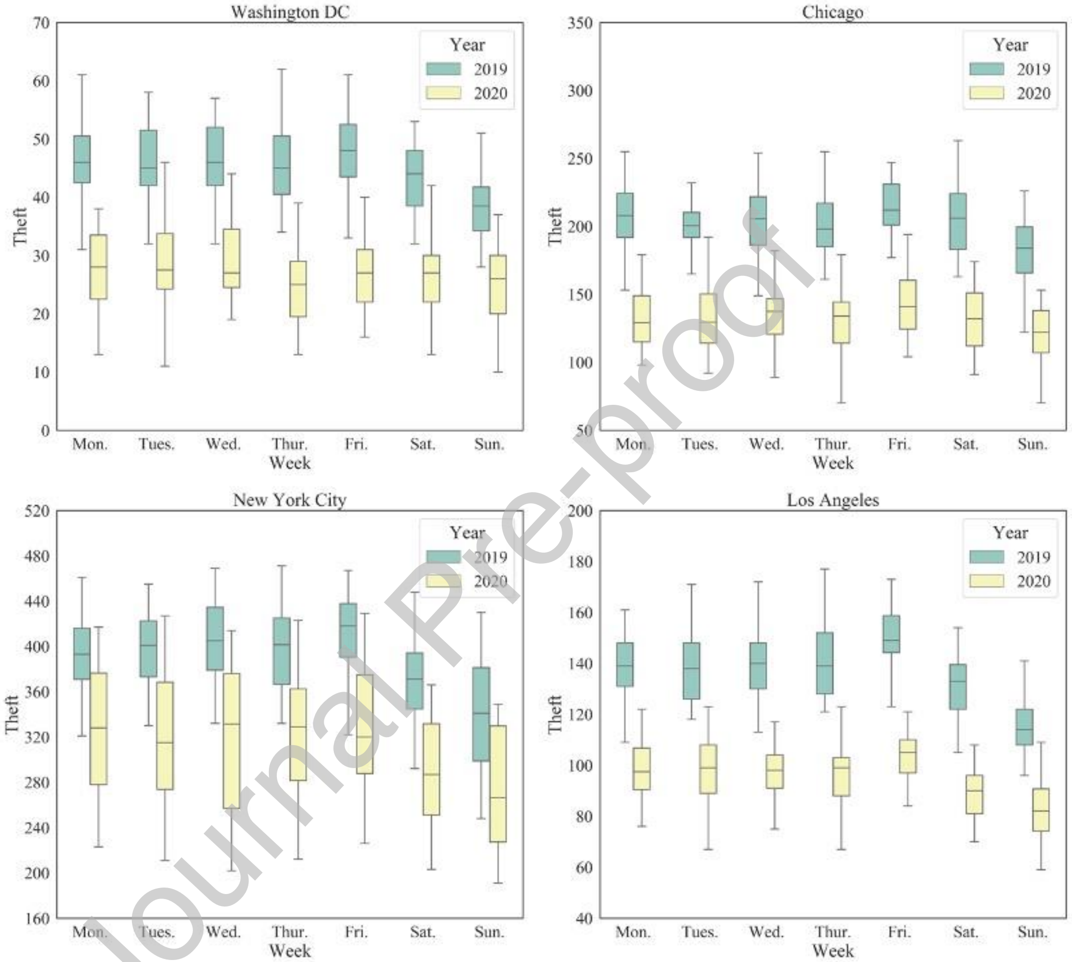

Figure 3. The distributions of daily theft incidents numbers before and during the COVID-19 pandemic in different cities of the U.S. The green boxplots represent the daily theft incidents numbers before the pandemic (2019) while the yellow ones are during the pandemic (2020).

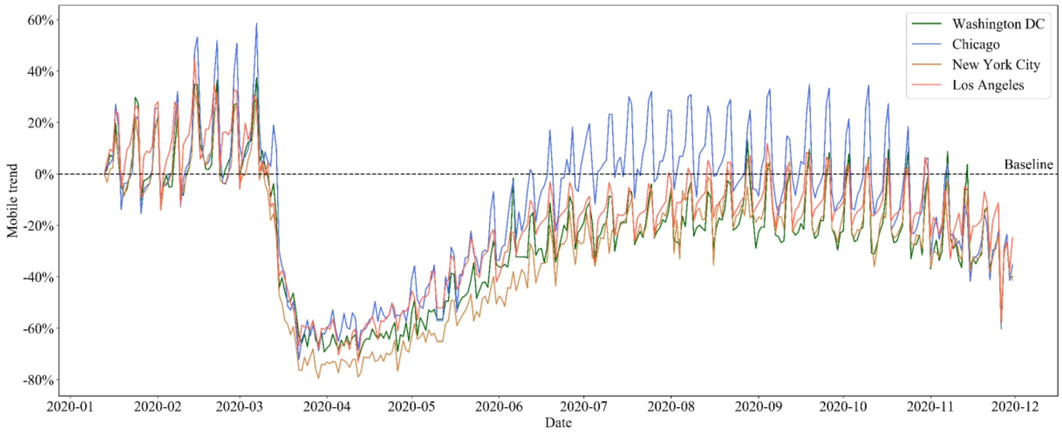

Figure 4. The mobile trends before and during the COVID-19 pandemic in different cities of the U.S. Green, blue, brown and red line represent the new confirmed cases of COVID-19 in DC, Chicago, New York City and Los Angeles, respectively. Black dotted line represents the baseline of the mobile trends.

3.2 The Granger causality between COVID-19 pandemic and crime

To answer the question whether COVID-19 pandemic influences crimes, Granger causality test is conducted. First, stationarity of the time series is examined and ensured to avoid spurious regression. Then, the optimal lags are selected by vector autoregressive models and the results are shown in Table 4. After that, the Granger causality test between crime incidents numbers and new confirmed cases of COVID-19 in different cities is implemented, and the results are also shown in Table 4.

As shown in Table 4, several Granger causal relationships are found in these US cities. For example, new confirmed cases of COVID-19 Granger cause the theft incidents numbers in DC since the *p* value is lower than 0.05, which means that the relationship is significant. In the US cities, both the number of crime incidents and the new confirmed cases of COVID-19 changed considerably during the periods studied in this paper. So, it is relatively easier to study their statistical laws.

Table 4. The results of the Granger causality test between crime incidents numbers and new cases of COVID-19 in different cities.

City	Null Hypothesis	Optimal lags	AIC	F-Statistic	*p* value	Conclusion
Washington DC	COVID-19 pandemic does not Granger cause theft	7	16.168	3.746	**0.025***	Refuse
	COVID-19 pandemic does not Granger cause fraud	7	17.830	3.627	**0.028***	Refuse
	COVID-19 pandemic does not Granger cause assault	7	13.407	1.591	0.139	Accept
	COVID-19 pandemic does not Granger cause robbery	7	10.908	0.458	0.864	Accept
	COVID-19 pandemic does not Granger cause burglary	7	12.534	0.416	0.892	Accept
Chicago	COVID-19 pandemic does not Granger cause theft	10	21.740	2.407	**0.010***	Refuse
	COVID-19 pandemic does not Granger cause fraud	10	20.450	0.849	0.582	Accept
	COVID-19 pandemic does not Granger cause assault	9	20.651	3.737	**0.000***	Refuse
	COVID-19 pandemic does not Granger cause robbery	14	19.952	0.794	0.684	Accept
	COVID-19 pandemic does not Granger cause burglary	9	23.574	1.413	0.183	Accept
Los Angeles	COVID-19 pandemic does not Granger cause theft	7	23.670	1.371	0.218	Accept
	COVID-19 pandemic does not Granger cause fraud	7	22.406	2.068	**0.040***	Refuse
	COVID-19 pandemic does not Granger cause assault	8	24.265	1.508	0.155	Accept
	COVID-19 pandemic does not Granger cause robbery	7	21.753	2.293	**0.028***	Refuse
	COVID-19 pandemic does not Granger cause burglary	7	24.280	0.438	0.878	Accept
New York City	COVID-19 pandemic does not Granger cause theft	7	24.813	2.361	**0.032***	Refuse
	COVID-19 pandemic does not Granger cause fraud	7	20.920	0.852	0.546	Accept
	COVID-19 pandemic does not Granger cause assault	7	24.679	2.142	**0.042***	Refuse
	COVID-19 pandemic does not Granger cause robbery	7	22.520	2.130	**0.043***	Refuse
	COVID-19 pandemic does not Granger cause burglary	3	25.063	0.348	0.791	Accept

* Denotes a significance level lower than 0.05.

3.3 Crime prediction with conducting new cases of COVID-19

Based on the results of Granger causality test between crime incidents numbers and new cases of COVID-19, several Granger causal relationships are confirmed in the US cities, which motivates us to conduct the new cases of COVID-19 into crime prediction in these cities, and to examine whether the new cases of COVID-19 can improve the prediction performance of the daily crime incidents numbers. In this study, new cases of COVID-19 that only with significant Granger causality is tried to improve the crime prediction by LSTM models. The results of ADF test are shown in Table A1, and the parameters of LSTM models are shown in Table A2.

Fig. 5 shows the predictions of daily crime incidents numbers for two weeks in different cities of the US, in which both the results with and without conducting COVID-19 are shown. As shown in Fig. 5, the predicted values are approximately consistent with the observations. Moreover, the prediction results with and without conducting the new COVID-19 cases are quite close to each other.

Figure 5. Prediction of daily crime incidents numbers in different cities of the US. Black lines and points represent the real daily crime incidents numbers (Observed), red lines and points represent the predicted crime incidents numbers without conducting the COVID-19 cases (Predicted), while green lines and points represent the predictions with conducting the COVID-19 cases (Predicted with COVID-19).

In order to quantitatively evaluate and compare the prediction results, the indices RMSE and PRMSE are calculated and their values are shown in Table 5. Indicated by them, models conducting COVID-19 cases performs slightly better than those without the cases. This demonstrates that involving new cases of COVID-19 as a variable can improve the performance of crime prediction in terms of some specific types of crime.

Table 5. Evaluations for the crime predictions.

City	Crime	Crime prediction		Crime prediction (with COVID-19)	
		RMSE	PRMSE	RMSE	PRMSE
Washington, DC	Theft	8.175	27.38%	7.220	24.18%
	Fraud	6.441	26.84%	5.268	25.13%
Chicago	Theft	15.434	13.03%	9.713	8.20%
	Assault	6.853	15.47%	6.171	13.93%
Los Angeles	Fraud	4.427	38.50%	4.154	36.12%
	Robbery	5.399	25.98%	4.973	23.93%
New York City	Theft	84.877	27.40%	73.383	23.69%
	Assault	28.924	17.33%	28.150	16.92%
	Robbery	9.479	23.96%	8.927	22.56%

4 Conclusion

This study investigates the impact of COVID-19 pandemic on crimes in four large cities (namely Washington DC, Chicago, New York City and Los Angeles). The differences of crime incidents numbers before and during the pandemic are investigated, and the Granger causal relationships between crime incidents numbers and new cases of COVID-19 are examined. Then, significant correlations between COVID-19 and crimes are used to improve the crime prediction performance.

Overall, the result shows that crime is indeed impacted by the COVID-19 pandemic, but it varies in different cities and also with different crime types. Most types of crimes have seen fewer incidents numbers during the pandemic than before. For example, theft numbers decrease significantly in these cities. Moreover, in three of the US cities, theft numbers are proved Granger caused by the new cases of COVID-19. For some other crime types and cities, similar results are reported. This may be partially explained by the Routine Activity theory and opportunity theory that people may prefer to stay at home to avoid being infected with COVID-19 during the pandemic, giving fewer chances for crimes.

Although providing some new insights on the relation between the COVID-19 pandemic and crimes, our work comes with some limitations. For the reason of data limitation, the research area only involves American cities. For future work, more data of other large cities around the world are recommended to use for investigating the impact of the pandemic on crime. This may be more helpful to compare the results between different countries all through the world. Apart from that, some other variables are expected to be extracted from the COVID-19 pandemic as new indices (such as variations of travel frequency, individual income etc.), which may be useful to explore deeper relations and laws between pandemic and crime as well as to improve the crime prediction performance.

Appendix A

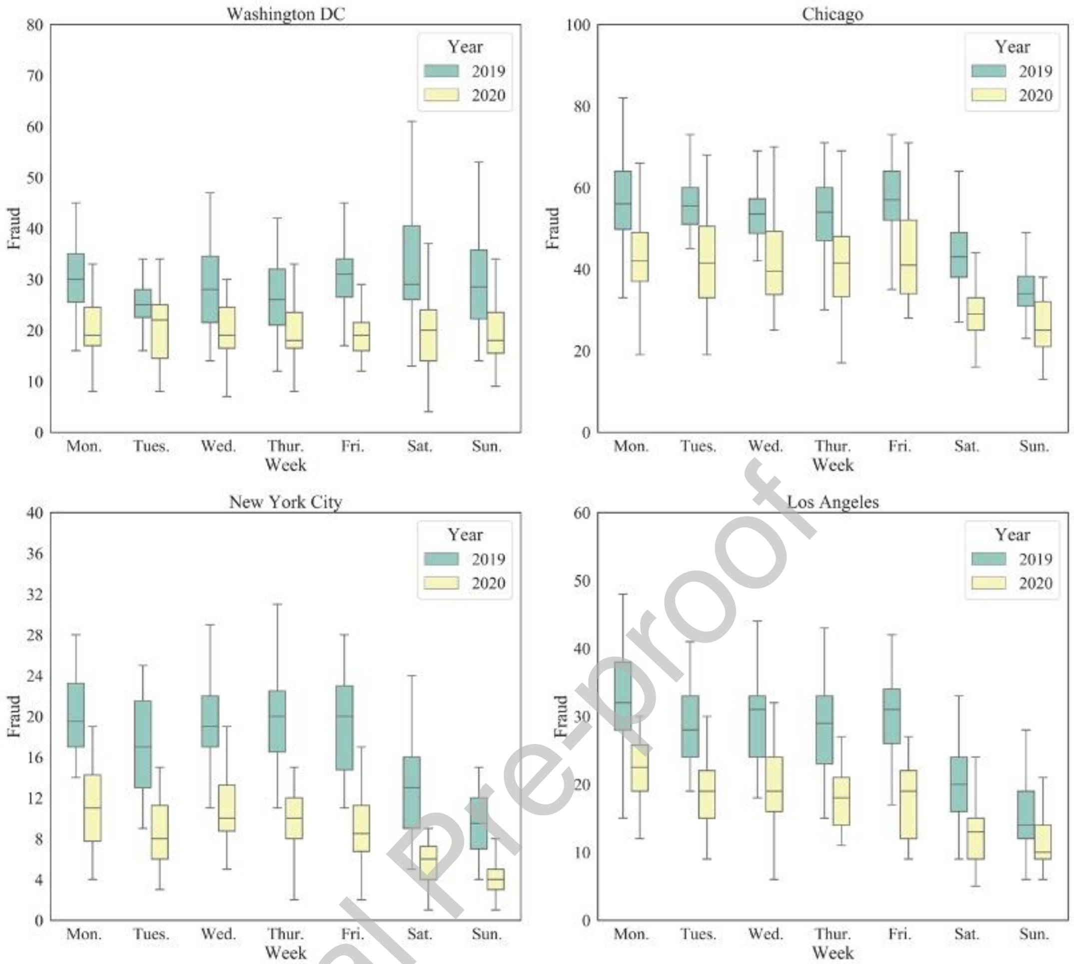

Figure A1. The distributions of daily fraud incidents numbers before and during the COVID-19 pandemic in different cities of the U.S. The green boxplots represent the daily fraud incidents numbers before the pandemic (2019) while the yellow ones are during the pandemic (2020).

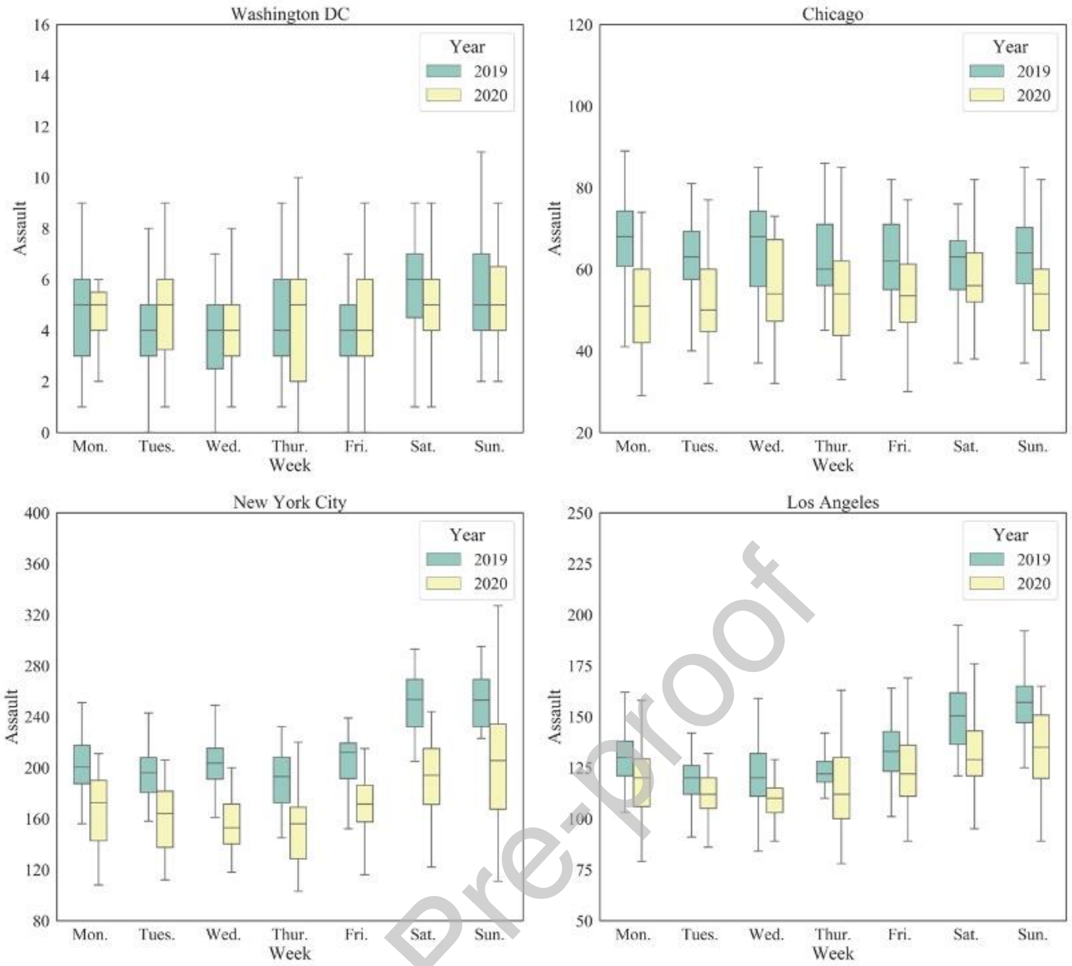

Figure A2. The distributions of daily assault incidents numbers before and during the COVID-19 pandemic in different cities of the U.S. The green boxplots represent the daily assault incidents numbers before the pandemic (2019) while the yellow ones are during the pandemic (2020).

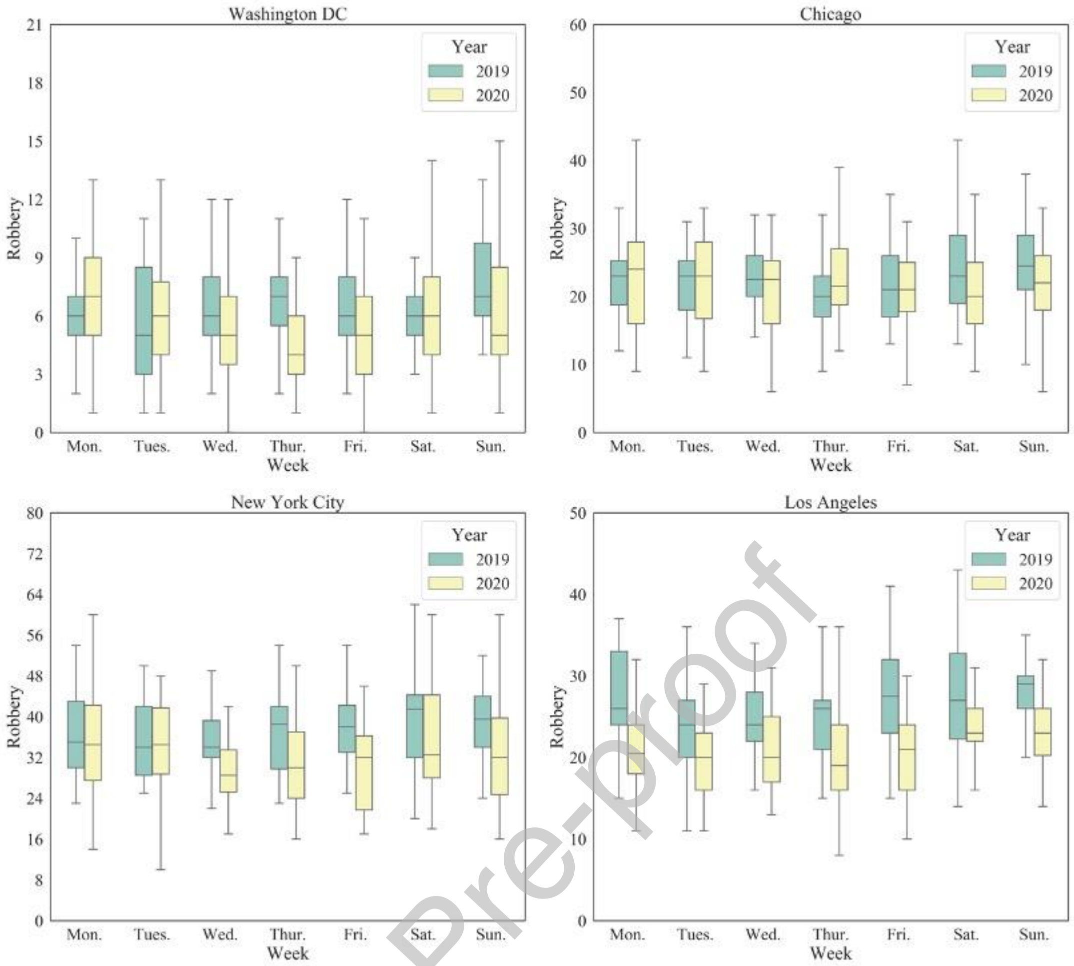

Figure A3. The distributions of daily robbery incidents numbers before and during the COVID-19 pandemic in different cities of the U.S. The green boxplots represent the daily robbery incidents numbers before the pandemic (2019) while the yellow ones are during the pandemic (2020).

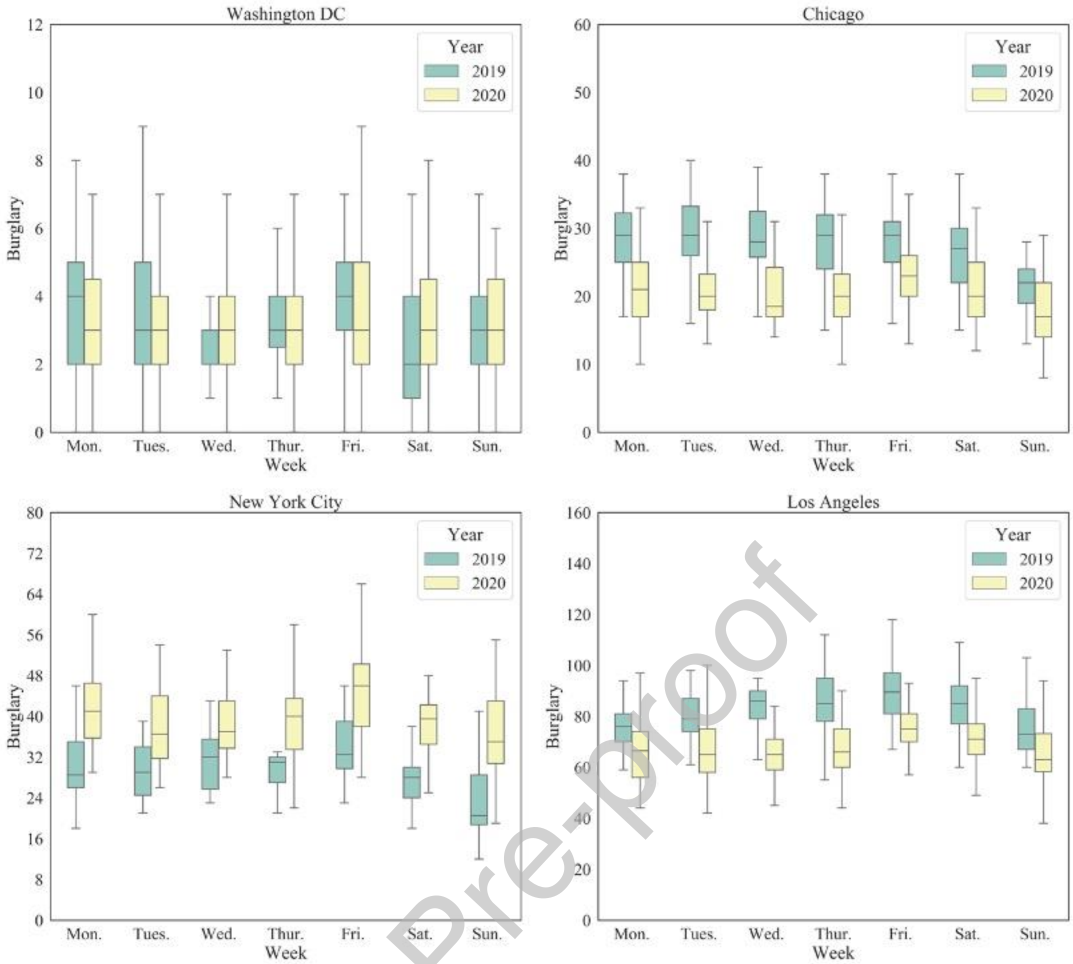

Figure A4. The distributions of daily burglary incidents numbers before and during the COVID-19 pandemic in different cities of the U.S. The green boxplots represent the daily burglary incidents numbers before the pandemic (2019) while the yellow ones are during the pandemic (2020).

Table A1. The results of the Augmented Dickey-Fuller test for different crime incidents numbers during the COVID-19 pandemic in different cities.

City	Variable	ADF Statistic	*p* value	Critical Values 1% Level	5% Level	10% Level	Conclusion
Washington, DC	Theft	-3.523	0.007*	-3.432	-2.862	-2.567	Stationary
	Fraud	-4.552	0.000*	-3.432	-2.862	-2.567	Stationary
Chicago	Theft	-2.025	0.276	-3.432	-2.862	-2.567	Non-stationary
	D(Theft)	-15.279	0.000*	-3.432	-2.862	-2.567	Stationary
	Assault	-3.810	0.003*	-3.432	-2.862	-2.567	Stationary
Los Angeles	Fraud	-4.279	0.000*	-3.432	-2.862	-2.567	Stationary
	Robbery	-3.972	0.002*	-3.432	-2.862	-2.567	Stationary
New York City	Theft	-4.600	0.000*	-3.432	-2.862	-2.567	Stationary
	Assault	-4.423	0.000*	-3.432	-2.862	-2.567	Stationary
	Robbery	-3.516	0.001*	-3.432	-2.862	-2.567	Stationary

* Denotes a significance level lower than 0.05. D (Theft) represents the first difference of the theft time series.

Table A2. The parameters of LSTM models.

City	Variable	Batch Size	Epochs	Hidden Neurons
Washington, DC	Theft	1	50	3
	Fraud	1	50	1
Chicago	Theft	1	20	2
	Assault	1	10	6
Los Angeles	Fraud	1	30	8
	Robbery	1	30	1
New York City	Theft	1	10	3
	Assault	1	10	3
	Robbery	1	20	7

References

[1] F.M. Khan, R. Gupta, ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India, Journal of Safety Science Resilience 1 (2020) 12-18.

[2] F.M. Khan, A. Kumar, H. Puppala, G. Kumar, R. Gupta, Projecting the Criticality of COVID-19 Transmission in India Using GIS and Machine Learning Methods, Journal of Safety Science Resilience 2 (2021) 50-62.

[3] Y. Zeng, X. Guo, Q. Deng, S. Luo, H. Zhang, Forecasting of COVID-19: spread with dynamic transmission rate, Journal of Safety Science Resilience 1 (2020) 91-96.

[4] A. Abu-Rayash, I. Dincer, Analysis of mobility trends during the COVID-19 coronavirus pandemic: Exploring the impacts on global aviation and travel in selected cities, Energy research social science 68 (2020) 101693.

[5] M.R. Fatmi, COVID-19 impact on urban mobility, Journal of Urban Management 9 (2020) 270-275.

[6] A.F. Siegenfeld, Y. Bar-Yam, The impact of travel and timing in eliminating COVID-19, Communications Physics 3 (2020) 1-8.

[7] V. Venkatesh, Impacts of COVID-19: A research agenda to support people in their fight, International journal of information management 55 (2020) 102197.

[8] J. Daniel, Education and the COVID-19 pandemic, Prospects 49 (2020) 91-96.

[9] L. Wen, M. Sheng, B. Sharp, The impact of COVID-19 on changes in community mobility and variation in transport modes, New Zealand Economic Papers (2021) 1-8.

[10] G. Caperna, M. Colagrossi, A. Geraci, G. Mazzarella, (Technical report, 2020).

[11] A. Şahin, M. Tasci, J.J.E.c. Yan, The unemployment cost of COVID-19: How high and how long?, (2020).

[12] E. Workie, J. Mackolil, J. Nyika, S. Ramadas, Deciphering the impact of COVID-19 pandemic on food security, agriculture, and livelihoods: A review of the evidence from developing countries, Current Research in Environmental Sustainability 2 (2020) 100014.

[13] S. Shayegh, M. Malpede, Staying home saves lives, really!, EIEE 2 (2020) 1-12.

[14] G.M. Campedelli, A. Aziani, S. Favarin, Exploring the immediate effects of COVID-19 containment policies on crime: an empirical analysis of the short-term aftermath in Los Angeles, American Journal of Criminal Justice 46 (2020) 1-24.

[15] M. Felson, S. Jiang, Y. Xu, Routine activity effects of the Covid-19 pandemic on burglary in Detroit, March, 2020, Crime Science 9 (2020) 1-7.

[16] J.R.B. De la Miyar, L. Hoehn-Velasco, A. Silverio-Murillo, Druglords don't stay at home: COVID-19 pandemic and crime patterns in Mexico City, Journal of Criminal Justice 72 (2021) 101745.

[17] M.P. Ashby, Initial evidence on the relationship between the coronavirus pandemic and crime in the United States, Crime Science 9 (2020) 1-16.

[18] R. Rosenfeld, R. Fornango, The impact of economic conditions on robbery and property crime: The role of consumer sentiment, Criminology 45 (2007) 735-769.

[19] R. Harries, Modelling and predicting recorded property crime trends in England and Wales—a retrospective, International Journal of Forecasting 19 (2003) 557-566.

[20] D.L. Yearwood, G. Koinis, Revisiting property crime and economic conditions: An exploratory study to identify predictive indicators beyond unemployment rates, The Social Science Journal 48 (2011) 145-158.

[21] M. De Nadai, Y. Xu, E. Letouzé, M.C. González, B. Lepri, Socio-economic, built environment, and mobility conditions associated with crime: a study of multiple cities, Scientific reports 10 (2020) 1-12.
[22] S.K. Rumi, K. Deng, F.D. Salim, Crime event prediction with dynamic features, EPJ Data Science 7 (2018) 43.
[23] L.G. Alves, H.V. Ribeiro, F.A. Rodrigues, Crime prediction through urban metrics and statistical learning, Physica A: Statistical Mechanics its Applications 505 (2018) 435-443.
[24] J. Wang, J. Hu, S. Shen, J. Zhuang, S. Ni, Crime risk analysis through big data algorithm with urban metrics, Physica A: Statistical Mechanics its Applications 545 (2020) 123627.
[25] Z. He, L. Tao, Z. Xie, C. Xu, Discovering spatial interaction patterns of near repeat crime by spatial association rules mining, Scientific reports 10 (2020) 1-11.
[26] S. Towers, S. Chen, A. Malik, D. Ebert, Factors influencing temporal patterns in crime in a large American city: A predictive analytics perspective, PLoS one 13 (2018) e0205151.
[27] E.G. Cohn, The prediction of police calls for service: The influence of weather and temporal variables on rape and domestic violence, Journal of Environmental Psychology 13 (1993) 71-83.
[28] X. Han, X. Hu, H. Wu, B. Shen, J. Wu, Risk Prediction of Theft Crimes in Urban Communities: An Integrated Model of LSTM and ST-GCN, IEEE Access 8 (2020) 217222-217230.
[29] K. Carbone-Lopez, J. Lauritsen, Seasonal variation in violent victimization: Opportunity and the annual rhythm of the school calendar, Journal of quantitative criminology 29 (2013) 399-422.
[30] L.G. Alves, E.K. Lenzi, R.S. Mendes, H.V. Ribeiro, Spatial correlations, clustering and percolation-like transitions in homicide crimes, EPL 111 (2015) 18002.
[31] Y. Jinghua, H.J.D.A. Miaomiao, K. Discovery, Study on Time Series Prediction of Theft Crime Based on LSTM Network, (2020) 1.
[32] J. Tiihonen, P. Halonen, L. Tiihonen, H. Kautiainen, M. Storvik, J. Callaway, The association of ambient temperature and violent crime, Scientific reports 7 (2017) 1-7.
[33] M. Ranson, Crime, weather, and climate change, Journal of environmental economics management 67 (2014) 274-302.
[34] S. Field, The effect of temperature on crime, The British journal of criminology 32 (1992) 340-351.
[35] X. Hu, J. Wu, P. Chen, T. Sun, D. Li, Impact of climate variability and change on crime rates in Tangshan, China, Science of the Total Environment 609 (2017) 1041-1048.
[36] B. Shen, X. Hu, H. Wu, Impacts of climate variations on crime rates in Beijing, China, Science of the Total Environment 725 (2020) 138190.
[37] E.G.J.T.B.J.o.C. Cohn, Weather and crime, 30 (1990) 51-64.
[38] R. Peet, The geography of crime: A political critique, The Professional Geographer 27 (1975) 277-280.
[39] E. Beqiraj, S. Fedeli, L. Giuriato, Policy tolerance of economic crime? An empirical analysis of the effect of counterfeiting on Italian trade, European Journal of Political Economy 65 (2020) 101933.
[40] F.K. Bappee, L.M. Petry, A. Soares, S. Matwin, Analyzing the impact of foursquare and streetlight data with human demographics on future crime prediction, arXiv preprint arXiv:.07516 2 (2020) 1-15.
[41] M. Chodos, B. Sarani, A. Sparks, B. Bruns, S. Gupta, C.P. Michetti, J. Crane, E. Hall, C.T. Trankiem, C. Abouassaly, Impact of COVID-19 pandemic on injury prevalence and pattern in the Washington, DC Metropolitan Region: a multicenter study by the American College of Surgeons Committee on Trauma, Washington, DC, Trauma Surgery & Acute Care Open 6 (2021) e000659.

[42] L.R. Bullinger, J.B. Carr, A. Packham, (National Bureau of Economic Research, 2020).
[43] S. Zangari, D.T. Hill, A.T. Charette, J.E. Mirowsky, Air quality changes in New York City during the COVID-19 pandemic, Science of the Total Environment 742 (2020) 140496.
[44] A.I. Friedson, D. McNichols, J.J. Sabia, D. Dave, (National Bureau of Economic Research, 2020).
[45] C.W. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica: journal of the Econometric Society (1969) 424-438.
[46] D.A. Dickey, W.A. Fuller, Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica: journal of the Econometric Society (1981) 1057-1072.
[47] H. Akaike, A new look at the statistical model identification, IEEE transactions on automatic control 19 (1974) 716-723.
[48] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (1997) 1735-1780.
[49] X. Ma, Z. Tao, Y. Wang, H. Yu, Y. Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, Transportation Research Part C: Emerging Technologies 54 (2015) 187-197.
[50] X. Hu, D. Li, H. Huang, S. Shen, E. Bou-Zeid, Modeling and sensitivity analysis of transport and deposition of radionuclides from the Fukushima Dai-ichi accident, Atmospheric Chemistry and Physics 14 (2014) 11065-11092.
[51] A.E. Nivette, R. Zahnow, R. Aguilar, A. Ahven, S. Amram, B. Ariel, M.J.A. Burbano, R. Astolfi, D. Baier, H.-M. Bark, A global analysis of the impact of COVID-19 stay-at-home restrictions on crime, Nature Human Behaviour 5 (2021) 1-10.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: