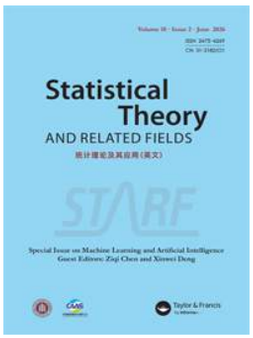

Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

Optimal prediction threshold and spatial characteristics of burglary crime across cities

Rui Wang , Yaofeng Zhang & Jinling Yao

To cite this article: Rui Wang , Yaofeng Zhang & Jinling Yao (2026) Optimal prediction threshold and spatial characteristics of burglary crime across cities, Statistical Theory and Related Fields, 10:2, 308-340, DOI: 10.1080/24754269.2026.2665854

To link to this article: https://doi.org/10.1080/24754269.2026.2665854

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 07 May 2026.

Submit your article to this journal

Article views: 68

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tstf20

STATISTICAL THEORY AND RELATED FIELDS
2026, VOL. 10, NO. 2, 308–340
https://doi.org/10.1080/24754269.2026.2665854

Taylor & Francis
by informa

OPEN ACCESS Check for updates

Optimal prediction threshold and spatial characteristics of burglary crime across cities

Rui Wang[a*], Yaofeng Zhang[a,b,c*] and Jinling Yao[d]

[a]Interdisciplinary Research Institute in New Finance and Economics, Hubei University of Economics, Wuhan, People's Republic of China; [b]Hubei Research Center for Digital Government Construction, Hubei University of Economics, Wuhan, People's Republic of China; [c]Hubei Centre for Data and Analysis, Hubei University of Economics, Wuhan, People's Republic of China; [d]School of Finance, Hubei University of Economics, Wuhan, People's Republic of China

ABSTRACT

Burglary, as a prevalent and detrimental crime type, poses a major threat to public safety and property security. Accurate prediction of burglary occurrence is therefore critical. Although deep learning has achieved notable progress in crime prediction, the influence of varying urban spatial characteristics on predictive performance and resource efficiency remains underexplored. This study analyzes burglary prediction in eight representative cities from China, the United States, and Canada, uniformly employing the model based on convolutional neural networks and long short-term memory networks (CNN–LSTM) under a consistent spatio–temporal scale. A comprehensive evaluation framework based on precision, hit rate, prediction accuracy index (PAI), and prediction effectiveness index (PEI) was established, focussing on the validity of PEI and variations in optimal prediction thresholds across cities. Furthermore, standard deviation ellipses, Moran index, and kernel density analysis were applied to quantify spatial characteristics and explore their associations with PEI and resource allocation. The results indicate that cities with more concentrated spatial distributions and stronger spatial autocorrelation exhibit superior predictive efficiency and resource utilization. This study enriches the analytical scope of crime prediction efficiency and supports a shift from 'accuracy–oriented' to 'efficiency–oriented' modelling for intelligent allocation of urban public safety resources.

ARTICLE HISTORY
Received 13 November 2025
Revised 10 March 2026
Accepted 23 April 2026

KEYWORDS
Crime prediction; urban spatial structure; PEI; CNN-LSTM; spatial analysis

1. Introduction

Criminal activities pose multidimensional threats to modern society, not only endangering individual lives and the safety of property but also exerting systemic, adverse effects on social stability and economic development. Among these, burglary, as a high-frequency crime, constitutes a persistent threat to public order, citizens' property rights, and privacy rights owing to its spatial diffusivity, the continuous nature of offenses, and the diversity of targets. Therefore, the exploration of criminal hotspot distribution patterns and the accurate prediction of

CONTACT Jinling Yao yyjinling@126.com School of Finance, Hubei University of Economics, No. 8, Yangqiao Lake Avenue, Canglong Island Development Zone, Jiangxia District, Wuhan, 430205, People's Republic of China

*These authors contributed equally to this work and should be considered co-first authors.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

burglary incidents have become core issues in public safety governance (He et al., 2023; Jin et al., 2020). Through the construction of a scientifically grounded crime prediction model, law enforcement agencies can achieve spatial optimization and dynamic, precise allocation of police resources, formulate data-driven crime prevention strategies, and implement proactive interventions prior to the occurrence of crimes, thereby significantly reducing crime rates and systematically enhancing the effectiveness of public safety governance.

With the growing demand for refined urban safety governance and the rapid advancement of big data technology, crime prediction research has rapidly evolved from classical statistical approaches to deep learning techniques. Early studies primarily relied on traditional statistical approaches, including the near-repeat model (Michael et al., 2003; Polvi et al., 1991), kernel density estimation (Chainey & Ratcliffe, 2005), and self-exciting point models (Mohler et al., 2017; Short et al., 2014). These approaches primarily utilized the spatio-temporal location of crime incidents and auxiliary information such as crime type to estimate future high-risk areas (Gu et al., 2021). Subsequently, machine learning algorithms, such as random forests (Alves et al., 2018; Liu et al., 2018) and support vector machines (Kianmehr & Alhajj, 2006, 2008), were widely adopted owing to their nonlinear modelling capabilities and feature-fusion advantages. These algorithms are capable of integrating multi-dimensional data, including crime distribution, socioeconomic characteristics, and road structure, to enhance predictive performance. In recent years, deep learning models have become the mainstream focus of research due to their superior ability to capture nonlinear temporal patterns and complex spatial dependencies (Gu et al., 2021). Foundational architectures such as long short-term memory (LSTM) networks (Shen et al., 2019) and convolutional neural networks (CNNs) (Matereke et al., 2021) have been widely applied. For instance, Li et al. (2025) proposed a hybrid CNN-LSTM model integrated with data binning preprocessing, significantly enhancing prediction accuracy in Philadelphia. Similarly, Maheshwari et al. (2025) demonstrated that ConvLSTM outperforms traditional methods in forecasting crime hotspots across both urban and rural contexts, effectively revealing distinct spatiotemporal dynamics.

More recently, the field has witnessed a shift towards advanced spatiotemporal modelling paradigms, particularly graph neural networks (GNNs) and transformers. GNN-based approaches excel at modelling non-Euclidean spatial correlations inherent in road networks and urban communities. Roshankar and Keyvanpour (2023) constructed a spatio-temporal graph neural network (STGNN) using Chicago crime data, which captures dynamic spatial proximity through graph structures, achieving an accuracy of 82% and outperforming baseline models. Furthermore, He et al. (2023) developed a road-weight STGNN (RW-STGNN) that explicitly incorporates road network weights to refine spatial dependency learning. Parallel to GNNs, transformer architectures have been introduced to address long-range temporal dependencies. Wang et al. (2024) proposed a multi-dimensional spatio-temporal fusion transformer (MDSTFT), combining CNNs, attention mechanisms, and Transformer layers to improve multivariate time-series forecasting precision. These models demonstrate that integrating complex topological structures (e.g., road networks, interaction graphs) and attention mechanisms significantly boosts predictive performance.

Although significant progress has been achieved in model architecture innovation and multi-source data integration, most existing studies have concentrated on optimizing model performance and algorithmic innovation with the aim of improving predictive accuracy. However, notable shortcomings remain in terms of the applicability of crime prediction across diverse urban contexts, the rationality of threshold selection, and the systematic

influence of urban spatial characteristics on predictive efficiency. Specific issues include the following: first, cities are constructed in accordance with their topography and terrain, exhibiting spatial characteristics and structures such as monocentric, bicentric, and polycentric configurations. When combined with variations in road network density and functional zoning, these factors result in substantial differences in the applicability and predictive performance of identical models across diverse cities. Existing research has primarily been confined to single-city studies (Esquivel et al., 2020; Zhao et al., 2023), with limited systematic cross-city comparisons or structured syntheses. Second, the threshold for crime prediction–namely, the determination of the number of prediction hotspots–has often relied on empirical judgment (Rummens & Hardyns, 2020; Rummens et al., 2017). No scientific and universally applicable evaluation criteria have yet been established, which complicates the balance between resource allocation and prediction accuracy. The number of prediction hotspots directly influences both the precision and the practical value of prediction outcomes. An excessive number of hotspots may lead to resource inefficiency and reduced accuracy, whereas too few may overlook high-risk areas and thereby weaken the practical utility of predictions. Third, although the relationship between the spatial clustering of crime and predictive performance has been examined in some studies (Chainey et al., 2008), systematic analyses and cross-city comparative investigations remain lacking regarding how urban spatial characteristics affect predictive efficiency, particularly with respect to balancing model resource allocation and predictive effectiveness.

To address the aforementioned research gap, eight cities in China, the United States, and Canada (Wuhan, Los Angeles, Chicago, New York, Philadelphia, Minneapolis, San Francisco, and Montréal) were selected as the subjects of investigation. With a focus on spatial feature differences in crime prediction, the study aims to investigate how the spatial structural characteristics of different cities affect predictive effectiveness and resource efficiency. It is noteworthy that this study does not aim to directly enhance prediction accuracy. Instead, a unified prediction modelling framework is employed to systematically compare the relationship between spatial characteristics and predictive performance across multiple cities. Building upon the work of Zhang et al. (2023), the rationality and applicability of the prediction effectiveness index (PEI) in evaluating predictive effectiveness are further validated. Moreover, differences in optimal prediction thresholds across cities are examined based on the PEI metric, thereby revealing strategies to maximize the utilization of predictive resources under varying spatial characteristics. In doing so, the study highlights the adaptability and cost-efficiency of prediction models in terms of resource allocation.

To ensure the scientific rigour of cross-city comparisons and the comparability of results, a unified framework was employed across all methods, utilizing a crime prediction model that integrates convolutional neural networks (CNN) (Lecun et al., 2002) and long short-term memory (LSTM) networks (Hochreiter & Schmidhuber, 1997). Systematic modelling and prediction experiments were performed on eight cities using identical spatio-temporal scales and standardized input structures. To examine the influence of urban spatial structure on prediction performance, three dimensions were considered: spatial distribution concentration, spatial autocorrelation, and urban spatial structure type. Standard deviation ellipses (Lefever, 1926), Moran index (Moran, 1950), and kernel density analysis (Parzen, 1962; Rosenblatt, 1956) were applied to quantitatively analyze the crime distribution characteristics of each city. The relationships with the PEI indicator and optimal prediction thresholds were then systematically analyzed, experiences for improving prediction efficiency under different spatial characteristics were summarized, and targeted recommendations for resource

allocation and model optimization were proposed. By integrating deep learning prediction models with spatial statistical analysis, understanding differences in crime prediction effectiveness from a spatial structure perspective was enhanced, thereby advancing theoretical development regarding mechanism explanations and model adaptability in cross-city crime prediction research.

Compared with prior research, the primary contributions of this study are reflected in the following aspects. (1) From a cross-city comparative perspective, the applicability of a unified prediction method is systematically evaluated across different countries and cities, thereby addressing limitations in previous cross-city studies. (2) The spatial structural characteristics of crime are incorporated into the prediction efficiency analysis framework for the first time, and their mechanisms of influence on prediction threshold selection and PEI performance are explored. (3) The study introduces PEI as a comprehensive evaluation metric that integrates both predictive effectiveness and resource efficiency, promoting a paradigm shift in crime prediction from 'accuracy-oriented' to 'efficiency-oriented', and providing a reference pathway for optimizing resource allocation in public safety management. (4) The optimal threshold characteristics for crime prediction across different types of cities are summarized, offering practical guidance for police resource deployment in urban crime prevention.

2. A spatio-temporal prediction and analysis framework for crime based on CNN-LSTM

To enable cross-city comparison of crime prediction performance, a unified CNN-LSTM spatiotemporal prediction and analysis framework was developed. The framework encompasses the complete process, from data preparation and spatiotemporal input construction to model training and evaluation. In addition, spatial statistical methods are integrated to analyze the relationship between prediction outcomes and urban spatial characteristics, thereby establishing the foundation for subsequent cross-city comparative analysis.

2.1. Research sample and data preparation

Burglary crime data from eight cities – Wuhan, Los Angeles, Chicago, New York, Philadelphia, Minneapolis, San Francisco, and Montréal – were selected as the research sample. With the exception of the Wuhan dataset, which was obtained from the Big Data Practical Application Center of the Public Security Bureau, all data were collected from publicly available sources on the Kaggle platform. Each dataset contains essential attributes, including case type (type), incident time (time), longitude (lon), and latitude (lat).

Let C be the set of cities under study. For each city $c \in C$, the original data set is denoted as

$$D_c = \left\{ \left(t_i, \lambda_i, \varphi_i, \tau_i \right) \right\}_{i=1}^{N_c}. \tag{1}$$

Among these, t_i represents the timestamp of the case, (λ_i, φ_i) represents the longitude and latitude coordinates of the case, τ_i represents the case type label, and N_c represents the total amount of raw data for the city. The basic information of the urban crime dataset is shown in Table 1.

To enhance data quality and satisfy the model's input format requirements, the original data were preprocessed, specifically through data selection and data cleaning.

Table 1. Basic information of the crime datasets.

City	Time range	Records	Data source
Wuhan	3 Jan 2015–9 Jun 2019	934,698	Big-Data 110 Police-Call Platform, Public Security Bureau
Los Angeles	1 Jan 2010–9 Sep 2017	1,584,316	https://www.kaggle.com/cityofLA/crimeinlosangeles
Chicago	1 Jan 2001–11 Nov 2018	6,747,040	https://www.kaggle.com/spirospolitis/chicagocrimes20012018november
New York	26 Sep 2013–31 Dec 2015	104,857	https://www.kaggle.com/adamschroeder/crimesnewyorkcity
Philadelphia	1 Jan 2006–23 Mar 2017	2,237,605	https://www.kaggle.com/mchirico/philadelphiacrimedata
Minneapolis	1 Jan 2010–1 Jul 2016	136,121	https://www.kaggle.com/mrisdal/minneapolisincidentscrime
San Francisco	6 Jan 2003–3 May 2015	878,049	https://www.kaggle.com/competitions/sf-crime/overview
Montréal	1 Jan 2015–20 Aug 2021	136,642	https://www.kaggle.com/datasets/stevieknox/Montreal-crime-data

(1) Data selection

Because different types of cases exhibit distinct characteristics, their occurrence patterns are influenced by multiple factors. Burglary represents one of the most prevalent crime types, characterized by high incidence, broad geographical distribution, and a propensity to occur in series. It poses a significant threat to public safety and infringes on citizens' property and privacy rights. Therefore, burglary-related criminal incidents were selected as the focus of this study. A theft-related label set $T_{\text{burglary}} = \{\text{'', 'BURGLARY'}\}$ was defined, and fuzzy matching was applied to extract all data $D_c' = \{(t_i, \lambda_i, \varphi_i, \tau_i) \in D_c \mid \tau_i \in T_{\text{burglary}}\}$ associated with theft-related cases.

(2) Data cleaning

The objective of data cleaning is to eliminate redundant and invalid records from the burglary dataset D_c' obtained through preliminary screening, thereby ensuring the accuracy of subsequent spatio-temporal modelling. Data cleaning procedures include deduplication, deletion of missing values, and exclusion of data outside the study area.

Data deduplication: When two or more records $(t_i, \lambda_i, \varphi_i, \tau_i)$ and $(t_j, \lambda_j, \varphi_j, \tau_j)$ satisfy $t_i = t_j$, $\lambda_i = \lambda_j$, $\varphi_i = \varphi_j$, $\tau_i = \tau_j$, they are regarded as duplicates of the same case, and only the unique record is retained.

Missing value removal: When a case record lacks information on time ($t_i = \emptyset$), longitude ($\lambda_i = \emptyset$), or latitude ($\varphi_i = \emptyset$), the record is deleted.

Exclusion of non-study area data: Let the longitude and latitude range of city c be $\lambda^c_{\min} \leqslant \lambda \leqslant \lambda^c_{\max}$, $\varphi^c_{\min} \leqslant \varphi \leqslant \varphi^c_{\max}$. All records outside this range are excluded.

Through the aforementioned data cleaning procedures, the final cleaned dataset D_c^* was obtained, ensuring that each record in D_c^* represents a valid and unique burglary case within the study area and provides high-quality input for subsequent modelling.

2.2. Spatio-temporal data processing and input construction

After the extraction and cleaning of raw crime data are completed, the data must be converted into spatio-temporal structured input suitable for direct processing by deep learning models. As crime events inherently possess both spatial and temporal attributes, the appropriate characterization of their spatial proximity and temporal sequence during data processing directly affects the predictive performance and generalization capacity of subsequent models. To this end, spatial gridding and event mapping, temporal discretization and sliding window configuration, and spatio-temporal tensor construction are sequentially performed during the spatio-temporal data processing phase. These procedures transform discrete crime records into standardized, model-ready multi-dimensional spatio-temporal data representations, thereby yielding high-dimensional tensor data that comply with the input requirements of hybrid deep learning architectures such as CNN-LSTM.

2.2.1. Spatial gridding and crime data mapping

To achieve spatial discretization of the study city c while preserving spatial location information and local statistical characteristics for subsequent modelling, an equal-scale gridding method is employed to partition the continuous geographic space into cells of uniform physical size. The crime data are then mapped to the corresponding cells, thereby enabling the spatial clustering of crime events.

The longitude range of the minimum bounding rectangle of the study area c is denoted as $[\lambda^c_{\min}, \lambda^c_{\max}]$ and the latitude range is denoted as $[\varphi^c_{\min}, \varphi^c_{\max}]$. Accordingly, its spatial set is defined as

$$A^c = \left\{(\lambda, \varphi) \mid \lambda^c_{\min} \leqslant \lambda \leqslant \lambda^c_{\max}, \varphi^c_{\min} \leqslant \varphi \leqslant \varphi^c_{\max}\right\} \tag{2}$$

where (λ, φ) denotes a geographic coordinate point. Because the differences in longitude and latitude cannot be directly equated to actual distances, the physical lengths of the east-west and north-south directions of the study area must be computed prior to grid division. Let $\mathrm{Dist}(\cdot, \cdot)$ denote the ellipsoidal distance function. The east-west and north-south distances of the study area are then given by

$$D^c_{\text{lon}} = \mathrm{Dist}((\lambda^c_{\min}, \varphi^c_m), (\lambda^c_{\max}, \varphi^c_m)), \tag{3}$$

$$D^c_{\text{lat}} = \mathrm{Dist}((\lambda^c_m, \varphi^c_{\min}), (\lambda^c_m, \varphi^c_{\max})), \tag{4}$$

where $\lambda^c_m = (\lambda^c_{\min} + \lambda^c_{\max})/2$ and $\varphi^c_m = (\varphi^c_{\min} + \varphi^c_{\max})/2$ denote the longitude and latitude of the centre of the study area, respectively.

To ensure that the actual area of the grid cells approximates the desired value, the target unit spatial dimensions $L_x \times L_y$ (in kilometers) are first specified. Based on this specification, the number of grid columns n^c_x and rows n^c_y within the study area is calculated as follows:

$$n^c_x = \left\lceil \frac{D^c_{\text{lon}}}{L_x} \right\rceil, \tag{5}$$

$$n^c_y = \left\lceil \frac{D^c_{\text{lat}}}{L_y} \right\rceil. \tag{6}$$

The span of each grid in terms of longitude and latitude is then obtained:

$$\Delta\lambda^c = \frac{\lambda^c_{\max} - \lambda^c_{\min}}{n^c_x}, \tag{7}$$

$$\Delta\varphi^c = \frac{\varphi^c_{\max} - \varphi^c_{\min}}{n^c_y}. \tag{8}$$

For the sth crime data point, given its longitude and latitude coordinates $(\lambda^c_s, \varphi^c_s)$, the column and row indices of the corresponding grid cell are determined as follows:

$$x^c_s = \left\lfloor \frac{\lambda^c_s - \lambda^c_{\min}}{\Delta\lambda^c} \right\rfloor + 1, \tag{9}$$

$$y^c_s = \left\lfloor \frac{\varphi^c_s - \varphi^c_{\min}}{\Delta\varphi^c} \right\rfloor + 1. \tag{10}$$

The corresponding grid number is subsequently derived:

$$g^c_s = \left(y^c_s - 1\right) \cdot n^c_x + x^c_s. \tag{11}$$

Table 2. Basic statistics of the processed crime dataset.[1]

City	Number of records after preprocessing	Maximum longitude	Minimum longitude	Maximum latitude	Minimum latitude	Total number of grids
Wuhan	65,514	115.0769	113.6925	31.3622	29.972	20,770
Los Angeles	114,385	−117.6463	−118.9517	34.8233	32.75	3648
Chicago	386,552	−87.524081	−87.940101	42.02304	41.644531	1505
New York	13,006	−73.7002	−74.2588	40.9176	40.4766	2352
Philadelphia	93,979	−74.955832	−75.280298	40.13796	39.867005	868
Minneapolis	13,006	−93.193859	−93.329127	45.05125	44.89015	198
San Francisco	36,754	−122.28148	−123.17383	37.929811	37.640314	2607
Montréal	43,002	−73.474295	−73.974157	45.70479	45.410076	1320

Thus, all cases assigned to the same grid number are regarded as belonging to the same spatial unit, thereby achieving spatial aggregation. The indicator function is defined as

$$\mathbf{1}\left\{\left(x_s^c, y_s^c\right) = (i,j)\right\} = \begin{cases} 1, & \text{if the crime event } s \text{ is located at grid } (i,j), \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

On this basis, the number of crimes within each grid is recorded on a daily basis, thereby forming a two-dimensional frequency matrix:

$$N_{i,j}^c(t) = \sum_{s=1}^{S_t} \mathbf{1}\left\{\left(x_s^c, y_s^c\right) = (i,j)\right\}, \tag{13}$$

where S_t denotes the total number of cases on day t. To facilitate model training, this matrix is further binarized as follows:

$$B_{i,j}^c(t) = \mathbf{1}\left\{N_{i,j}^c(t) > 0\right\}, \tag{14}$$

where $B_{i,j}^c(t)$ indicates whether a crime occurred in grid (i,j) on day t. A summary of the crime dataset for the eight cities after grid processing is presented in Table 2. It is worth noting that the total number of spatial grids for city c is given by $n_x^c \times n_y^c$. In this study, the target physical size of each grid cell is set to 1000 m × 1000 m (i.e., $L_x = L_y = 1$ km). This spatial resolution is chosen based on empirical practice in crime geography as well as the typical operational range of police patrols, thereby achieving a reasonable balance between capturing spatial heterogeneity in crime patterns and maintaining computational efficiency. In addition, a unified grid size is adopted across all cities to ensure the comparability of spatial units in cross-city prediction analysis. Using different grid resolutions for different cities could introduce additional structural differences in spatial aggregation, which would make it difficult to distinguish whether variations in prediction performance arise from urban spatial structure or from differences in spatial discretization. Therefore, a consistent grid size is employed to maintain methodological comparability across cities.

2.2.2. Time discretization and sliding window settings

The proper handling of the temporal dimension is crucial for accurately capturing evolving trends and latent patterns in criminal behaviour. The data are discretized on a daily scale,

[1] The latitude and longitude information for each city's outer rectangle comes from https://boundingbox.klokantech.com/

with the continuous sequence of criminal events divided into spatial distribution slices, each representing one day.

The occurrence of criminal events typically exhibits strong temporal dependence and lag effects. To effectively capture this dynamic property, a sliding time window mechanism is introduced, integrating spatial information from multiple consecutive historical time slices to enhance the model's ability to learn temporal dependencies. Specifically, a sliding time window with a fixed length of L days ($L = 7$, i.e., one week Shen et al., 2019) is defined as the model input sequence length. For each prediction, the model utilizes the continuous spatial distributions from the preceding L days as input to forecast the crime distribution on day $L+1$. This design offers several advantages: first, the sliding window mechanism effectively incorporates temporal dependencies without explicitly introducing lagged variables, thereby avoiding redundancy and potential information leakage. Second, the 7-day time span smooths daily fluctuations while capturing both immediate triggers and cumulative effects of criminal behaviour, thereby enhancing the model's responsiveness to short- and medium-term trends and achieving a balance between prediction stability and sensitivity.

At the stage of model input construction, the spatial crime binary matrix for each time window is unfolded by time step to generate a three-dimensional input sequence. Specifically, for the sth sample, the input is defined as

$$X^c(s) = \left[B^c(s), B^c(s+1), \ldots, B^c(s+L-1)\right] \in \{0,1\}^{L \times n_x^c \times n_y^c}. \tag{15}$$

The corresponding label is defined as the binary matrix representing the crime distribution for the subsequent $s+L$ days:

$$Y^c(s) = B^c(s+L) \in \{0,1\}^{n_x^c \times n_y^c}. \tag{16}$$

The entire dataset therefore consists of $N = T-L$ such samples, thereby enabling continuous time-series segmentation and the systematic construction of supervised learning data.

In summary, the implementation of time discretization and sliding windows not only ensures the temporal continuity of the input data and fully leverages historical information but also strengthens the model's capacity to capture the spatio-temporal evolutionary patterns of criminal behaviour from both theoretical and practical perspectives, thereby improving predictive accuracy and practical applicability.

2.2.3. Spatio-temporal tensor construction

To ensure compatibility with the CNN-LSTM-based spatio-temporal deep learning model, the input data are further transformed into a multi-dimensional tensor that incorporates both time steps and channel dimensions.

Specifically, each sample input $X^c(s)$ is reshaped into a four-dimensional tensor $X^c(s) \in \mathbb{R}^{L \times n_x^c \times n_y^c \times C}$, where the time window length is set to $L = 7$, the spatial dimensions $n_x^c \times n_y^c$ correspond to the grid partitioning, and C denotes the number of channels. The corresponding label tensor is represented as $Y^c(s) \in \mathbb{R}^{n_x^c \times n_y^c}$. During training, the input tensor composed of all samples has dimensions $X^c(s) \in \mathbb{R}^{N \times L \times n_x^c \times n_y^c \times C}$, while the corresponding output labels are denoted as $Y^c(s) \in \mathbb{R}^{N \times n_x^c \times n_y^c}$, where N denotes the total number of samples. This tensor structure fully satisfies the model's input requirements for capturing dependencies across both spatial and temporal dimensions, thereby facilitating the joint learning of spatio-temporal associations.

2.3. Crime spatio-temporal prediction model based on CNN-LSTM

To enable dynamic spatio-temporal modelling of urban crime risks and to facilitate structural comparisons across cities, a CNN-LSTM-based crime spatio-temporal prediction model is constructed, which simultaneously extracts spatial patterns and captures temporal dependencies.

2.3.1. Model selection and explanation

LSTM is a deep learning architecture that enhances traditional recurrent neural networks (RNNs) by more effectively modelling long-term dependencies. By incorporating a gating mechanism, it selectively retains or discards information, thereby capturing long-term dependencies in time-series data more effectively. Since criminal behaviour generally exhibits pronounced time-series characteristics, with its distribution and evolution during specific periods strongly influenced by preceding events, the ability of LSTM to capture temporal dependencies allows it to filter noise and redundant information effectively while extracting key historical signals highly relevant to the prediction task. Therefore, it represents an ideal approach for crime prediction modelling (Mei & Li, 2019; Pimple et al., 2023).

However, crime data exhibit not only temporal correlations but also spatial dependencies. The use of LSTM alone neglects spatial proximity relationships, which makes it difficult to fully capture spatial structural information. To address this limitation, a CNN module is introduced at the input stage of the model. Through local convolution and pooling operations, features are extracted from the two-dimensional spatial distribution of each time segment. The spatial feature sequences from different time points are subsequently fed into the LSTM layer for temporal dependency modelling, thereby forming an end-to-end hybrid architecture that integrates CNN and LSTM. This architecture enables the simultaneous processing of spatial patterns and temporal evolution, balancing the extraction of local spatial structures with the capture of long-term temporal dependencies, thereby fulfilling the modelling requirements for the spatio-temporal coupling of crime data.

Traditional spatio-temporal statistical models, such as Bayesian spatio-temporal models and spatial econometric models, although advantageous in terms of statistical interpretability, typically rely on explicit prior assumptions and parameter specifications, thereby exhibiting notable limitations when applied to complex, nonlinear, and highly dynamic urban crime data (Hu et al., 2018; Law et al., 2014; Mohler et al., 2011; Ratcliffe, 2009). In contrast, the CNN-LSTM hybrid model adopted in this study provides substantial advantages. It does not rely on rigid prior assumptions and is capable of adaptively capturing multi-scale spatio-temporal patterns in high-dimensional, high-frequency, nonlinear, and dynamically evolving crime data. Moreover, the end-to-end training mechanism of CNN-LSTM facilitates the simultaneous optimization of spatial feature extraction and temporal dependency modelling, thereby improving both the expressive capacity and stability of predictions.

It should be emphasized that the maximization of model prediction accuracy is not the core objective of this study. Rather, the emphasis is placed on examining how spatial structure affects the relationship between prediction accuracy and resource utilization efficiency, thereby validating the applicability and explanatory power of the proposed PEI metric. To maintain methodological consistency, a unified CNN-LSTM architecture was employed to avoid structural discrepancies arising from multiple model comparisons, thereby allowing a clearer identification of prediction performance differences among cities attributable to spatial structural variations. These differences are reflected in multiple aspects, including

prediction accuracy, equity in resource distribution, and the outcomes of PEI indicator calculations, thereby providing robust evidence for cross-regional comparative analyses in urban public safety management and police resource allocation.

2.3.2. Model structure and training settings

The CNN-LSTM model constructed in this study integrates the two key components of spatio-temporal feature extraction and temporal dependency modelling within its structural design.

For each input crime tensor X_t ($t \in \{1, 2, \ldots, T\}$), a convolution operation is applied using a time-distributed two-dimensional convolution layer (TimeDistributed Conv2D):

$$F_t = \sigma\left(W_{\text{conv}} * X_t + b_{\text{conv}}\right), \tag{17}$$

where $*$ represents the convolution operation, W_{conv} and b_{conv} denote the convolution kernel weights and biases, respectively, and $\sigma(\cdot)$ refers to the ReLU activation function. Subsequently, the max pooling layer MaxPool$(\cdot)$ is employed to downsample the convolution features, thereby reducing dimensionality and computational burden:

$$P_t = \text{MaxPool}\left(F_t\right). \tag{18}$$

Next, the pooled feature P_t at each time step is flattened into a one-dimensional vector $f_t \in \mathbb{R}^d$, forming the input sequence $F = \{f_1, f_2, \ldots, f_T\}$, which is then fed into the LSTM layer. Temporal dependencies are captured by the LSTM unit through the following recursive formulations:

$$i_t = \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right), \tag{19}$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right), \tag{20}$$

$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o\right), \tag{21}$$

$$c_t = f_t c_{t-1} + i_t \alpha\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right), \tag{22}$$

$$h_t = o_t \alpha\left(c_t\right). \tag{23}$$

i, f and o represent the input gate, forget gate, and output gate, respectively. c denotes the cell state, which consists of the previous cell state value (regulated by the forget gate) and the current candidate memory state value (regulated by the input gate). W. represents the learnable parameter. The final output value h_t is produced through the combined action of the output gate and the current memory cell state.

The final hidden state of the model is mapped through a fully connected layer to estimate the crime probability of each spatial grid cell, thereby constituting the model output. With respect to model configuration, the binary cross-entropy loss function was employed to accommodate the binary classification characteristics of the prediction task. Regarding activation functions, the hidden layer utilized the tanh function to enhance nonlinear representational capacity, whereas the output layer employed the sigmoid function to produce probability-based predictions. The initial weights of the model were randomly initialized within the range $[0, 1]$, and the weights of each layer were iteratively optimized during training through the backpropagation algorithm to enhance predictive performance. For data partitioning, the dataset of each city was divided into training and test sets with a 7:3 ratio, respectively, used for model fitting and performance evaluation, thereby ensuring the

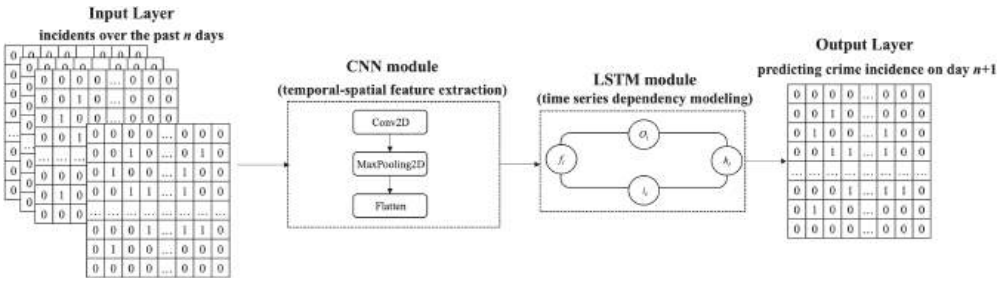

Figure 1. Crime spatio-temporal prediction model based on CNN-LSTM.

generalization capability of the model. The structural diagram of the CNN-LSTM-based spatio-temporal crime prediction model is presented in Figure 1.

2.4. Crime prediction effectiveness evaluation indicators

The accuracy of crime prediction models directly determines whether police operations can be guided with precision and effectiveness. In addition to precision, hit rate, the prediction accuracy index (PAI), and the prediction effectiveness index (PEI) was employed in this study to assess the effectiveness of crime prediction. The formulas corresponding to the evaluation indicators adopted in this study are provided as follows:

$$P = \frac{n}{m}, \tag{24}$$

where P denotes the precision rate, n represents the number of correctly predicted crime hotspots, and m indicates the number of hotspots predicted to experience crime.

$$R = \frac{n}{N}, \tag{25}$$

where R denotes the hit rate, n represents the number of correctly predicted crime hotspots, and N corresponds to the number of actual crime hotspots in the study area.

$$\mathrm{PAI} = \frac{\tilde{n}/\tilde{N}}{a/A}, \tag{26}$$

where PAI denotes the prediction accuracy index, $\tilde{n}$ represents the number of correctly predicted crime incidents, $\tilde{N}$ indicates the number of actual crime incidents in the study area, and $\tilde{n}/\tilde{N}$ reflects the case hit rate rather than the hotspot hit rate. Furthermore, a refers to the area of hotspots predicted to experience incidents, and A corresponds to the total area of the study region.

$$\mathrm{PEI} = \frac{2n}{m+N} \cdot \frac{1}{k_s}. \tag{27}$$

Among these, n is defined as the number of correctly predicted crime hotspots, m is defined as the number of hotspots predicted to experience crime, N is defined as the actual number of hotspots where crime occurred in the study area, and k_s denotes the multiple of the unit grid area relative to the standard grid area ($S^* = 500 \times 500\ \mathrm{m}^2$), thereby indicating the influence of the unit grid size on crime prediction effectiveness. By explicitly incorporating the

grid-area factor through k_s, the PEI indicator accounts for the influence of spatial discretization on prediction results, which helps ensure that evaluation outcomes remain comparable under different spatial aggregation settings. A higher PEI value indicates a closer alignment between predicted and actual crime occurrence patterns, thereby reflecting improved prediction accuracy and effectiveness. PEI balances the trade-off between hit rate and precision, which is influenced by the number of predicted hotspots – where an increase in hit rate results in a decrease in precision – while simultaneously reflecting the effect of unit grid size on prediction effectiveness (Zhang et al., 2023). Moreover, PEI enables reasonable control over both the number of predicted crime hotspots and the unit grid area, thereby providing a more accurate reflection of their influence on prediction results. Therefore, PEI was selected as the primary evaluation indicator for crime prediction results in this study, as it not only balances the trade-off between hit rate and precision but also explicitly considers the influence of unit grid size on prediction effectiveness.

2.5. Framework for analyzing spatial characteristics of crime

To elucidate the mechanisms through which the urban spatial structure influences crime prediction performance, this study constructs an analytical framework of crime spatial characteristics from three core dimensions: spatial distribution concentration, spatial autocorrelation, and urban spatial structure type. The framework aims to quantify the spatial distribution patterns of urban crime and systematically uncover their intrinsic relationships with PEI and the model's optimal prediction threshold, thereby enriching the understanding of the mechanisms underlying performance differences in spatio-temporal crime prediction models.

The concentration of spatial distribution is considered to reflect the intensity and directionality of crime events within a city. The analysis employing standard deviation ellipses has been shown to effectively capture the spatial extent and dominant orientation of crime distribution, thereby facilitating the identification of spatial layout trends of crime hotspots. By comparing variations in the major and minor axes and azimuth angles of crime distribution ellipses across different cities, the spatial morphological differences in crime concentration are quantified, providing a critical basis for the construction of spatial structural variables in subsequent models.

Spatial autocorrelation is considered to reflect the degree of interdependence and clustering of criminal incidents within geographical space. The Moran index, serving as a global spatial autocorrelation indicator, is used to quantify the spatial clustering trend of crime across the entire region, thereby indicating whether criminal incidents exhibit significant spatial clustering or dispersion. Through the calculation and analysis of the Moran index, differences in the strength of spatial associations between crimes in different cities are revealed, further elucidating the similarities and differences in the efficiency of spatial information utilization during the model prediction process.

The identification of urban spatial structure types is conducted using non-parametric methods, such as kernel density analysis, to precisely depict dense areas of crime hotspots and their spatial distribution patterns. Kernel density methods are employed to smooth crime event point data, generating continuous crime density distribution maps that intuitively reflect the spatial morphology and intensity of hotspots. By integrating kernel density results across different cities, the spatial structural differences in hotspot distributions are

Table 3. Comparison of functions of three spatial analysis methods.

Method	Analytical dimension	Main function	Output format	Application purpose
Standard Deviation Ellipse (SDE)	Global direction and dispersion	Describes the orientation and spatial spread of crime incidents	Elliptical diagram	Determines whether crime distribution exhibits directional patterns
Moran's I	Spatial autocorrelation	Measures the degree of spatial clustering of crime incidents	Numeric value (−1 to +1)	Identifies whether spatial clustering patterns exist
Kernel Density Estimation (KDE)	Local hotspot distribution	Displays local density variations of crime incident points within a geographic area	Heatmap	Reveals the location and intensity of crime hotspots

systematically compared, and the potential impact of spatial structural complexity on model performance is explored.

For clarity, Table 3 presents a concise comparison of the core functions of the three methods. By synergistically applying these methods, the spatial distribution characteristics of criminal incidents in the study area can be depicted from multiple dimensions, thereby establishing a methodological foundation for subsequent hotspot identification analyses.

3. Research results and analysis

3.1. Evaluation of crime prediction effectiveness at different hotspot thresholds

Using the CNN-LSTM model to predict crime in eight cities, the precision rate, hit rate, PAI and PEI were calculated separately. Formulas (24) to (27) indicate that the predicted number of hotspots (m), as a subjectively determined threshold, exerts a significant influence on the prediction results. Based on existing research (He et al., 2023; Rummens & Hardyns, 2020; Rummens et al., 2017), the relevant evaluation metrics were calculated using three methods for determining the hotspot threshold values. The three methods are as follows: (1) the number of predicted hotspots was set to 5% of the total number of grids, ensuring that the prediction results were broadly applicable in large-scale areas and avoiding excessive concentration in a few hotspot regions. The total number of grids is presented in Table 2; (2) the number of predicted hotspots was set to 100, providing a fixed standard that ensures consistency across different scenarios and facilitates comparative analysis; (3) the number of predicted hotspots was set to twice the average number of hotspots, where the latter represents the mean of crime hotspot counts over the seven days preceding the prediction date. This method allows flexible adjustment of the threshold to accommodate changes in actual conditions, thereby avoiding biases induced by excessively high or low thresholds. By applying these three hotspot threshold values, the prediction model's performance under varying conditions can be comprehensively evaluated, thereby ensuring the reliability and representativeness of the evaluation results. The precision rate, hit rate, PAI and PEI results under the three hotspot threshold settings are presented in Tables 4–6.

The above results indicate that varying threshold values produce different outcomes, with the precision rate, hit rate, and PAI being significantly influenced by the hotspot threshold. Therefore, the standard deviation and coefficient of variation of the evaluation indicators obtained under different threshold settings were calculated to assess their dispersion and relative stability, thereby facilitating the selection of reasonable evaluation metrics. The standard deviation is a commonly used statistical measure for quantifying the dispersion of values within a dataset, while the coefficient of variation (CV), defined as the ratio of the

Table 4. Evaluation results of the indicator of predicting the number of hotspots as 0.05 of the total number of grids.

	Wuhan	Los Angeles	Chicago	New York	Philadelphia	Minneapolis	San Francisco	Montréal
precision rate	0.023685	0.071806	0.138478	0.149044	0.128325	0.125529	0.10436	0.099791
hit rate	0.934117	0.354665	0.301689	0.438779	0.343512	0.149579	0.999807	0.470023
PAI	18.74498	7.162091	6.14242	9.081331	6.929101	3.006578	19.89736	9.697083
PEI	0.011634	0.030227	0.048266	0.055929	0.047471	0.034059	0.048397	0.041678

Table 5. Evaluation results of the indicator for predicting the number of hotspots taken as 100.

	Wuhan	Los Angeles	Chicago	New York	Philadelphia	Minneapolis	San Francisco	Montréal
precision rate	0.105031	0.082357	0.128338	0.15434	0.103464	0.073921	0.13253	0.084924
hit rate	0.39956	0.222765	0.368017	0.384862	0.62779	0.884418	0.969802	0.604014
PAI	86.49665	8.277012	5.674677	9.441409	5.498974	1.758006	25.36503	8.130639
PEI	0.041571	0.030333	0.04847	0.05535	0.04554	0.034922	0.059635	0.037866

Table 6. Evaluation results of the metric using twice the average number of hotspots for predicting the number of hotspots.

	Wuhan	Los Angeles	Chicago	New York	Philadelphia	Minneapolis	San Francisco	Montréal
precision rate	0.129364	0.083363	0.138478	0.15948	0.133954	0.119041	0.244748	0.131078
hit rate	0.285269	0.180572	0.301689	0.357747	0.29459	0.286392	0.613024	0.298948
PAI	107.8148	8.393627	6.14242	9.769411	7.272266	2.890868	49.53236	13.0395
PEI	0.044351	0.028737	0.048266	0.055403	0.046687	0.042182	0.088884	0.045805

standard deviation to the mean, reflects the relative variability of an indicator and is not affected by measurement units. In this study, these descriptive statistics are used to compare the relative stability of different evaluation indicators across prediction settings. Indicators with smaller standard deviations and coefficients of variation exhibit less fluctuation under different prediction thresholds and can therefore be considered relatively more stable.

The standard deviation and coefficient of variation of the evaluation indicators derived from the three threshold-setting methods across the eight cities were calculated separately, and the results are presented in Table 7. The results show that, compared with precision rate, hit rate, and PAI, the PEI indicator generally exhibits lower coefficients of variation across most cities. Specifically, among the eight cities examined, PEI shows the smallest coefficient of variation among the four indicators in seven cities, indicating that it remains relatively stable under different prediction coverage settings. Although the PEI coefficient of variation is not the smallest in San Francisco, the overall pattern across the majority of cities still demonstrates that PEI tends to fluctuate less under different threshold configurations.

3.2. Optimal prediction threshold

In crime prediction, the optimal prediction threshold is defined as the number of predicted hotspots that maximizes the PEI under a specified prediction model. This threshold represents the point at which the optimal balance between precision and recall is achieved. Exceeding the threshold implies that patrol and surveillance areas must be expanded, resulting in inefficient use of police resources and a reduction in prediction accuracy, whereas a predicted hotspot count below the threshold may lead to the omission of high-risk areas, consequently diminishing the effectiveness of preventive and control measures. Therefore,

Table 7. Standard deviation and coefficient of variation results of evaluation indicators obtained using different threshold value methods.

Wuhan	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.023685	0.105031	0.129364	0.05534335	0.08602667	0.64332783
hit rate	0.934117	0.39956	0.285269	0.34636622	0.53964867	0.64183651
PAI	18.744978	86.496651	107.814849	46.5084162	71.018826	0.65487447
PEI	0.011634	0.041571	0.044351	0.01813999	0.03251867	0.55783301
Los Angeles	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.071806	0.082357	0.083363	0.00640182	0.07917533	0.08085626
hit rate	0.354665	0.222765	0.180572	0.09081688	0.25266733	0.35943261
PAI	7.162091	8.277012	8.393627	0.67986871	7.94424333	0.08558005
PEI	0.030227	0.030333	0.028737	0.00089243	0.02976567	0.02998175
Chicago	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.138478	0.128338	0.138478	0.00585433	0.135098	0.04333396
hit rate	0.301689	0.368017	0.301689	0.03829449	0.32379833	0.11826648
PAI	6.14242	5.674677	6.14242	0.27005155	5.98650567	0.04511005
PEI	0.048266	0.04847	0.048266	0.00011778	0.048334	0.00243678
New York	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.149044	0.15434	0.15948	0.00521819	0.154288	0.03382113
hit rate	0.438779	0.384862	0.357747	0.04124813	0.393796	0.10474493
PAI	9.081331	9.441409	9.769411	0.34416458	9.430717	0.036494
PEI	0.055929	0.05535	0.055403	0.00032008	0.05556067	0.005761
Philadelphia	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.128325	0.103464	0.133954	0.01622444	0.12191433	0.13308067
hit rate	0.343512	0.62779	0.29459	0.17992109	0.421964	0.42638966
PAI	6.929101	5.498974	7.272266	0.94053088	6.56678033	0.14322557
PEI	0.047471	0.04554	0.046687	0.00097117	0.046566	0.02085577
Minneapolis	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.125529	0.073921	0.119041	0.02811078	0.10616367	0.26478716
hit rate	0.149579	0.884418	0.286392	0.39079859	0.44012967	0.88791696
PAI	3.006578	1.758006	2.890868	0.68989095	2.55181733	0.2703528
PEI	0.034059	0.034922	0.042182	0.0044616	0.03705433	0.1204071
San Francisco	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.10436	0.13253	0.244748	0.0742689	0.160546	0.46260201
hit rate	0.999807	0.969802	0.613024	0.21517122	0.86087767	0.24994402
PAI	19.89736	25.365025	49.53236	15.7701652	31.5982483	0.49908353
PEI	0.048397	0.059635	0.088884	0.02090053	0.06563867	0.31841801
Montréal	Predict 5% of the total grids	Predict 100	Predict 2∗avg hotspot	Standard deviation	Mean	Coefficient of variation
precision rate	0.099791	0.084924	0.131078	0.02355878	0.10526433	0.22380588
hit rate	0.470023	0.604014	0.298948	0.1529082	0.45766167	0.33410751
PAI	9.697083	8.130639	13.0395	2.50740011	10.2890723	0.24369545
PEI	0.041678	0.037866	0.045805	0.00397054	0.041783	0.09502768

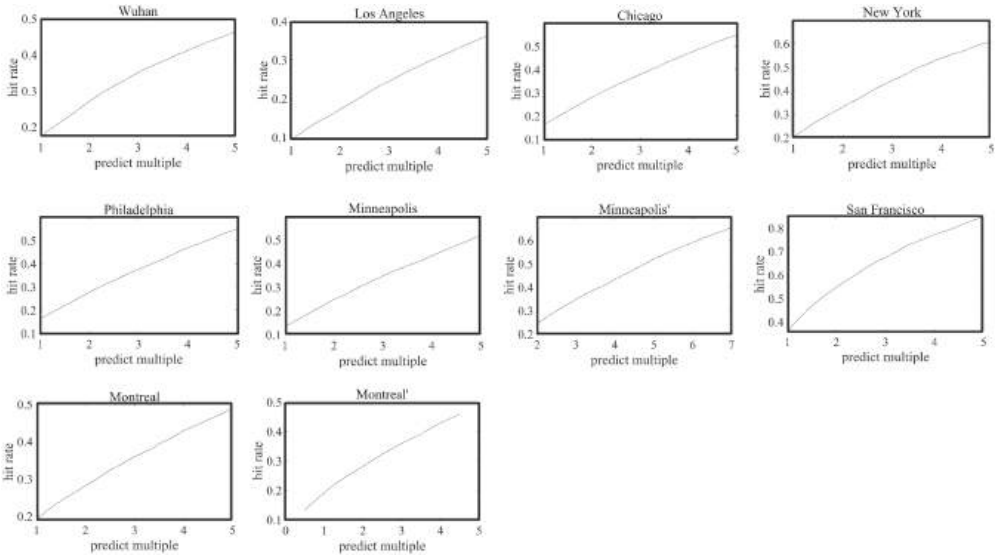

Figure 2. Hit rate.

the determination of the optimal prediction threshold is critical for enhancing prediction accuracy, practicality, and efficiency in resource utilization.

Zhang et al. (2023) indicates that PEI exhibits a 'rise-then-fall' trend with respect to changes in the number of predicted hotspots and proposes a dynamic-threshold-based prediction method: first, machine learning is employed to predict the number of crime hotspots N^* expected on the following day; second, the prediction multiplier k^* that maximizes PEI is determined experimentally; finally, the predicted number of hotspots is calculated as $m = N^* \times k^*$. To simplify calculations and analysis, N^* is replaced in this study by the average number of actual crime hotspots observed during the t days preceding the prediction day.

$$\bar{N} = \frac{1}{t}\sum_{i=1}^{t} N_i. \tag{28}$$

Set the prediction multiple set $k \in \{1 + 0.1n \mid n \in [0, 40]\}$, and different prediction multiples k are iterated on the test set. The performance of PEI and other evaluation metrics is then calculated, allowing the optimal prediction multiple k^* to be determined and the corresponding optimal number of predicted hotspots, i.e., the optimal prediction threshold:

$$m = \bar{N} \times k^*. \tag{29}$$

Thus, the problem of determining the optimal prediction threshold is equivalent to identifying the optimal prediction multiple. Given the weekly periodic temporal pattern of burglary noted by Shen et al. (2019), $t = 7$ was adopted in this study, meaning that the average number of actual crime hotspots over the seven days preceding the prediction date was used as $\bar{N}$. The mean values of the four evaluation metrics were subsequently calculated across different prediction multiples, and the results are presented in Figures 2–5.

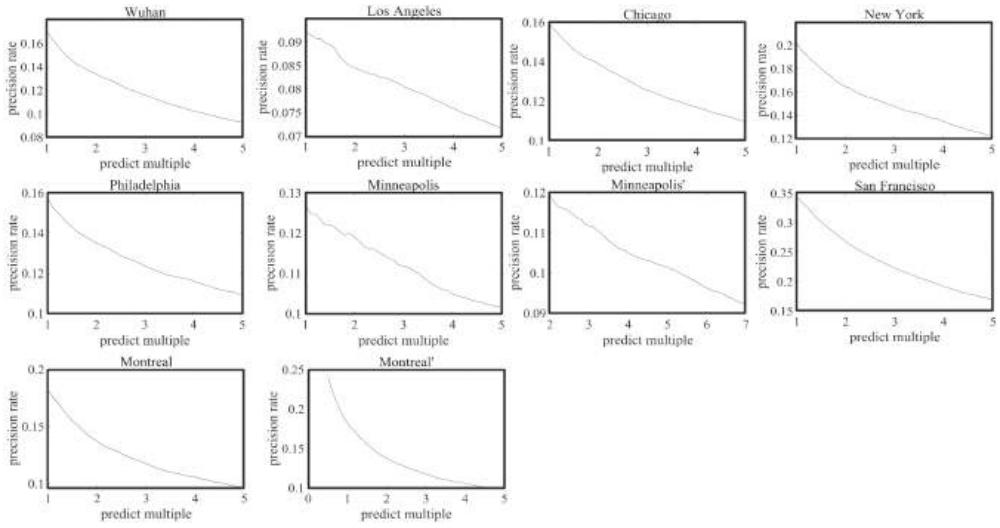

Figure 3. Precision rate.

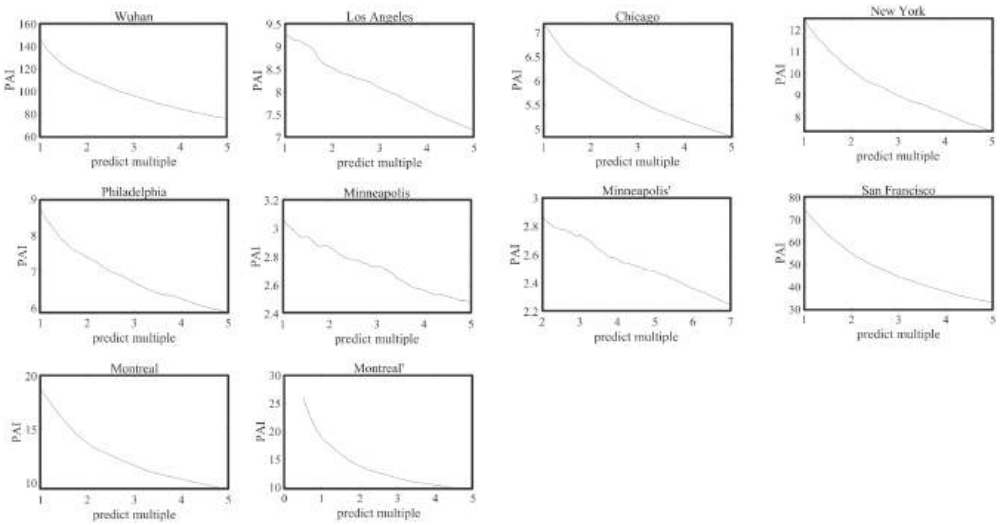

Figure 4. PAI.

As shown in the above figures, the hit rate increases monotonically as the prediction multiple increases, whereas the precision rate and PAI decrease monotonically. The PEI initially rises and then declines as the prediction multiple increases, forming a clear unimodal (inverted-*U*) relationship with the prediction multiple. This pattern reflects the trade-off between prediction coverage and resource efficiency: as the prediction multiple increases, more spatial grids are included in the prediction area, which improves the hit rate but gradually reduces the precision rate. From a resource allocation perspective, the marginal contribution of newly added prediction grids tends to decrease as the coverage expands, since these additional grids generally have lower crime occurrence probabilities. Consequently, the prediction efficiency first increases and then decreases, resulting in the observed unimodal

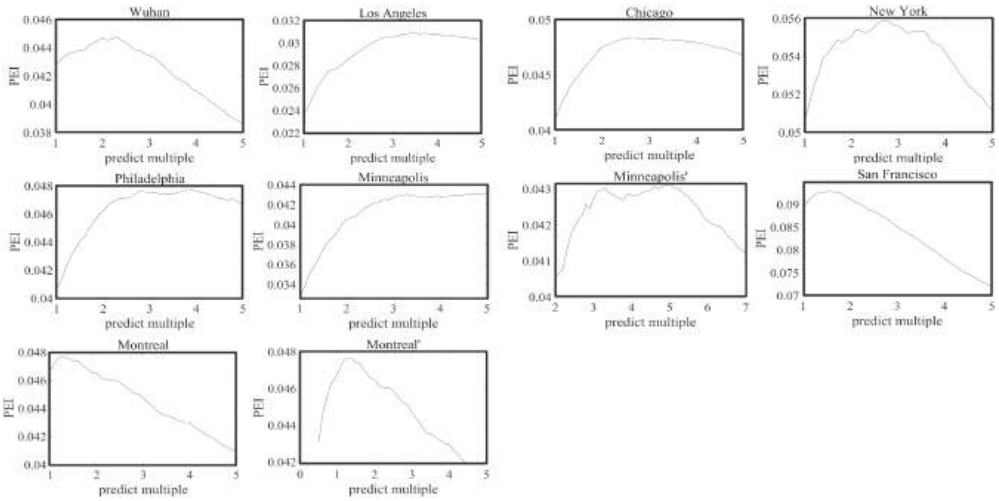

Figure 5. PEI.

pattern of PEI. In the analyses of Minneapolis and Montréal, however, this trend in PEI was not clearly observed when the prediction multiple was initially set to $k \in \{1 + 0.1n \mid n \in [0, 40]\}$, which may be attributed to the characteristics of crime distribution or underlying data patterns in these two cities. For example, the average number of crime hotspots in Minneapolis is relatively small, making the PEI curve less sensitive to changes in prediction coverage within the initial range. Consequently, the prediction multiple ranges were appropriately adjusted: for Minneapolis, because the downward trend at higher prediction multiples was not evident, the range was revised to $k \in \{2 + 0.1n \mid n \in [0, 40]\}$; for Montréal, because the upward trend at lower prediction multiples was not evident, the range was revised to $k \in \{0.5 + 0.1n \mid n \in [0, 40]\}$. The adjusted prediction results are presented in the figures titled Minneapolis' and Montréal', respectively. After this adjustment, the relationship between PEI and the prediction multiple was more clearly observed. Furthermore, the consistent unimodal relationship between PEI and the prediction multiple across cities indicates that PEI can reliably identify an optimal prediction coverage level that balances hit rate and precision. This structural characteristic provides additional evidence of the robustness of PEI as an evaluation indicator, complementing the stability analysis based on the coefficient-of-variation comparison presented in Table 7.

Those figures and the above arguments demonstrate that each city possesses an optimal prediction multiple that maximizes PEI, thereby producing the corresponding optimal prediction threshold. The set of prediction multiples ranking in the top 10 for PEI is defined as the 'optimal prediction multiple range', offering more flexible guidance for practical applications. To facilitate further investigation of the relationship between spatial characteristics and PEI, Table 8 reports the optimal prediction multiple k^* along with its corresponding optimal prediction multiple range for each city.

In practical applications, crime prediction models must be designed to account for the limitations of law enforcement resources. Determining the optimal prediction threshold facilitates an appropriate balance between prediction accuracy and hit rate, while also providing a foundation for evidence-based allocation of police resources and targeted prevention

Table 8. Optimal prediction multiple k^* and optimal prediction multiple range for cities.

City	Optimal prediction multiple k^*	Optimal prediction multiple range
Wuhan	2.3	1.7–2.6
Los Angeles	3.4	3.4–3.8
Chicago	2.7	2.4–3.3
New York	2.7	2.1; 2.4–3.1; 3.5
Philadelphia	3.8	2.8–3.1; 3.6–4.1
Minneapolis	5	3; 4.5–5.3
San Francisco	1.6	1.1–2.0
Montréal	1.3	1.0–1.9

and control in high-risk areas. The optimal prediction multiples (k^*) and the corresponding optimal prediction multiple ranges presented in Table 8 can serve as a reference for the development of tailored prevention and control strategies across different cities under varying crime distribution patterns and resource constraints. Furthermore, the proposed research framework and methodology can also be applied to hotspot prediction in other domains, such as healthcare resource allocation, traffic congestion management, and environmental monitoring, thereby providing theoretical support and methodological guidance for resource optimization and decision-making in these fields.

3.3. Relationship between spatial characteristics and PEI

Crime data from eight cities in 2015 were selected for spatial feature analysis. As the 2015 crime data for San Francisco were incomplete, the 2014 crime data were therefore used in the study. The spatial indicators used in this study are intended to capture the overall spatial organization of crime within each city rather than short-term temporal fluctuations. Urban crime spatial structures are generally relatively stable over medium time scales, while daily crime events tend to exhibit substantial randomness. Therefore, annual aggregated crime data were used to estimate spatial characteristics in order to reduce noise caused by short-term fluctuations and to better represent the underlying spatial configuration of crime distribution. The spatial feature analysis was conducted using standard deviation ellipse analysis, Moran index calculation, and kernel density estimation. All analyses were performed using the ArcGIS platform. In the subsequent analysis, these spatial indicators are correlated with the average PEI values over the entire prediction period to evaluate how long-term spatial crime structures influence the overall prediction efficiency of the CNN–LSTM model.

3.3.1. Concentration of crime spatial distribution and PEI

The standard deviation ellipse is a visualization-based method for describing the spatial distribution characteristics of a study object using global spatial statistical techniques. Specifically, it is constructed around the mean spatial centre of geographic features, with the standard deviations along the X and Y axes calculated to define the ellipse's axes, thereby encompassing the feature distribution. This ellipse is used to assess whether the distribution of elements is elongated in a particular direction.

In this study, the standard deviation ellipse method was employed to visualize the spatial distribution of crime hotspots across different cities, allowing the quantification of the centre of gravity, the distribution range, and the orientation of crime hotspots. These spatial statistical indicators were represented by the centre (latitude and longitude coordinates), area, and

Table 9. Standard deviation ellipse analysis results.

City	Central longitude	Central latitude	Semi-major axis (km)	Semi-minor axis (km)	Rotation angle (degrees)	Ellipse area (km^2)	Ellipse grid area ratio (km^2)	Optimal PEI
Wuhan	114.3088	30.5841	18.903	16.1384	58.1049	958.384	46.1427	0.0447
Los Angeles	−118.382	34.0851	10.1279	23.3195	140.2372	741.9721	203.3915	0.0309
Chicago	−87.6684	41.8328	5.9412	15.7608	152.5265	294.1721	195.4632	0.0483
New York	−73.921	40.7247	9.1033	14.4233	42.3059	412.4899	175.3784	0.0559
Philadelphia	−75.1469	39.9996	10.2667	4.797	55.9154	154.7218	178.251	0.0477
Minneapolis	−93.2692	44.9668	3.7371	6.3476	155.0583	74.5238	376.3826	0.0431
San Francisco	−122.426	37.7701	3.8714	3.367	55.5747	40.9503	15.7078	0.0929
Montréal	−73.6095	45.5259	10.3589	6.9402	63.5208	225.8599	171.1059	0.0476

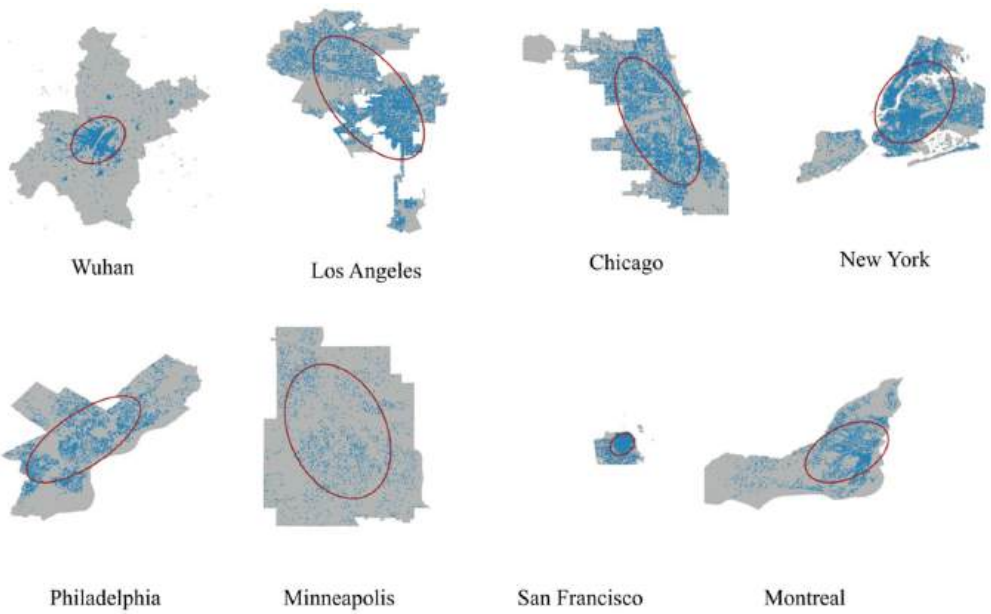

Figure 6. Standard deviation ellipse results.

azimuth of the standard deviation ellipse. To better capture the concentration and dispersion of crime hotspots and to control the influence of city size, an indicator termed the 'ellipse grid area ratio' was introduced, which is defined as the ratio of the ellipse area to the number of city grids. The ellipse grid area ratio serves as a measure of crime data dispersion within a city; higher values indicate greater dispersion of crime data, whereas lower values indicate greater concentration.

A standard deviation ellipse analysis was conducted on crime data from eight cities in 2015, which revealed the principal spatial distribution characteristics of crime in each city. The results are presented in Tables 9 and Figure 6. Specifically, the central longitude and central latitude represent the longitude and latitude of the ellipse centre, respectively. The semi-major axis and semi-minor axis correspond to the lengths of the major and minor axes of the standard deviation ellipse, respectively. The rotation angle denotes the directional angle of the ellipse, and the ellipse area refers to the area of the standard deviation ellipse.

The results indicate that regions with more concentrated crime distributions exhibit stronger predictive performance, as reflected in higher PEI values. For instance, San Francisco exhibited an elliptical grid area ratio of only 15.71, significantly lower than other cities, which indicates a highly concentrated spatial distribution of crime events. Its PEI value attained 0.0929, thereby achieving the highest predictive performance among all cities studied. By contrast, the elliptical grid area ratios for Los Angeles and Minneapolis were observed to be 203.39 and 376.38, respectively, suggesting a more dispersed spatial distribution of criminal incidents, accompanied by lower PEI values (0.0309 and 0.0431, respectively). This pattern suggests that spatial concentration may contribute to higher prediction efficiency under certain conditions. However, it is crucial to recognize the geometric limitations of the standard deviation ellipse. A larger SDE area might result from the long spatial distance between multiple isolated crime centres rather than a purely scattered distribution of individual incidents. Furthermore, spatial concentration is not the sole determinant of predictive performance. For instance, although crime in Wuhan is highly concentrated compared to the extremely dispersed pattern in Minneapolis, their PEI values are relatively close. This phenomenon indicates that predictive efficiency is a complex outcome influenced jointly by spatial concentration, the number of hotspots, and their temporal evolutionary stability. To overcome the limitations of SDE in accurately depicting complex polycentric structures, a peak dominance index (PDI) will be further introduced in Section 3.3.3.

Furthermore, the ratio of the major to minor axis of the standard deviation ellipse and the rotation angle revealed notable variations in the spatial structural characteristics of crime distributions across cities. Cities such as Wuhan, Los Angeles, and Chicago were characterized by ellipses with a substantially longer major axis relative to the minor axis, reflecting a pronounced directionality that may be influenced by urban development or geographical structures. In contrast, San Francisco's crime distribution approximated a circular form, indicative of isotropic spatial diffusion. These structural differences may further influence the effectiveness of models in capturing the spatiotemporal evolution of crime.

In summary, spatial concentration significantly impacts the resource efficiency of crime prediction models rather than merely their theoretical predictability. While CNNs are inherently capable of extracting local spatial features even in highly dispersed scenarios, scattered crime distributions require a broader predictive coverage area to capture isolated hotspots, inherently reducing resource utilization efficiency. More importantly, from a data characteristic perspective, extreme spatial dispersion leads to a significant decrease in daily crime counts within individual grids, exacerbating data sparsity. Such zero-inflated temporal sequences substantially reduce the signal-to-noise ratio for the LSTM module when attempting to learn stable temporal dependencies. Conversely, highly concentrated macro-structures provide dense, continuous spatiotemporal event sequences, creating optimal data conditions for the CNN-LSTM model to converge effectively and maximize predictive efficiency.

3.3.2. Spatial autocorrelation of crime and PEI

To analyze the spatial distribution characteristics and clustering trends of urban crime hotspots, Moran index is employed in this study as a key indicator for measuring spatial autocorrelation. Spatial autocorrelation is defined as the spatial dependency between variable values within a geographic space, that is, whether adjacent spatial units exhibit similar values. In crime prediction research, when a given area exhibits a high crime rate and its neighbouring areas also show high crime rates, it implies a positive spatial correlation, representing a pronounced spatial clustering pattern. Moran index is used to quantitatively measure this

Table 10. Moran index results and PEI.

City	Moran's I	*z* score	*p* value	Optimal PEI	Optimal prediction multiples k^*	Optimal prediction multiple range
Wuhan	0.844776	14.480338	0	0.044734	2.3	1.7–2.6
Los Angeles	0.230265	0.293967	0.768783	0.030881	3.4	3.4–3.8
Chicago	0.860679	7.863519	0	0.048342	2.7	2.4–3.3
New York	0.953846	9.58911	0	0.055891	2.7	2.1; 2.4–3.1; 3.5
Philadelphia	0.753693	5.957413	0	0.047732	3.8	2.8–3.1; 3.6–4.1
Minneapolis	0.575009	4.566834	0.000005	0.043149	5	3; 4.5–5.3
San Francisco	0.967744	20.84645	0	0.092853	1.6	1.1–2.0
Montréal	0.917917	12.41415	0	0.047637	1.3	1.0–1.9

spatial clustering pattern, and its formula is as follows:

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{30}$$

In this formula, n denotes the total number of urban grid cells; x_i denotes the number of criminal cases in the ith grid cell; $\bar{x}$ denotes the average number of criminal cases across all urban grid cells; and w_{ij} denotes the spatial weight matrix element, which describes the spatial proximity relationship between i and j.

By calculating the Moran index (Moran's I) of crime counts in each hotspot unit within a city, the presence of significant spatial clustering can be identified, thereby providing empirical evidence for analyzing the impact of spatial structure on the effectiveness of crime prediction. The Moran index results for each city were systematically analyzed in relation to the optimal values of the CNN-LSTM model's prediction performance (optimal PEI), optimal prediction multiplier, and optimal prediction multiplier range. The quantitative analysis results are summarized in Table 10.

Plotting the Moran index of each city against the optimal PEI in a scatter plot for analysis (Figure 7) indicates a positive association between the Moran index of crime spatial distribution and the optimal PEI. In particular, San Francisco (Moran's I $= 0.968$) and New York (Moran's I $= 0.954$) exhibit the strongest spatial clustering, and their predictive performance is the highest among the cities, with values of 0.0929 and 0.0559, respectively, ranking first and second. Conversely, Los Angeles has a Moran index of only 0.230 ($p > 0.76$), corresponding to the lowest predictive performance (PEI $= 0.0309$) and a nearly random spatial pattern. This trend suggests that the more spatially aggregated the distribution of crime, the more regular the spatiotemporal distribution of crime hotspots, facilitating the model's ability to capture their evolutionary trends and thereby achieve higher predictive accuracy.

The results presented in Table 10 further indicate that when the concentration of crime in a city is high, the optimal multiplier is typically smaller. For instance, San Francisco (1.6 times), Montréal (1.3 times), and Wuhan (2.3 times) exhibit high predictive performance at low multipliers, suggesting that their high-risk areas are relatively concentrated and have clear boundaries. In cities with weaker concentration, such as Los Angeles (3.4 times) and Minneapolis (5.0 times), higher multiplier thresholds are required to achieve optimal predictive performance, necessitating the expansion of model coverage to improve overall performance.

By comparing the optimal prediction multiplier ranges across cities, the robustness of predictive performance can thus be assessed, specifically regarding sensitivity to threshold parameter settings. For example, the optimal prediction multiplier ranges for San Francisco (1.1–2.0) and Montréal (1.0–1.9) are relatively concentrated, reflecting strong robustness of

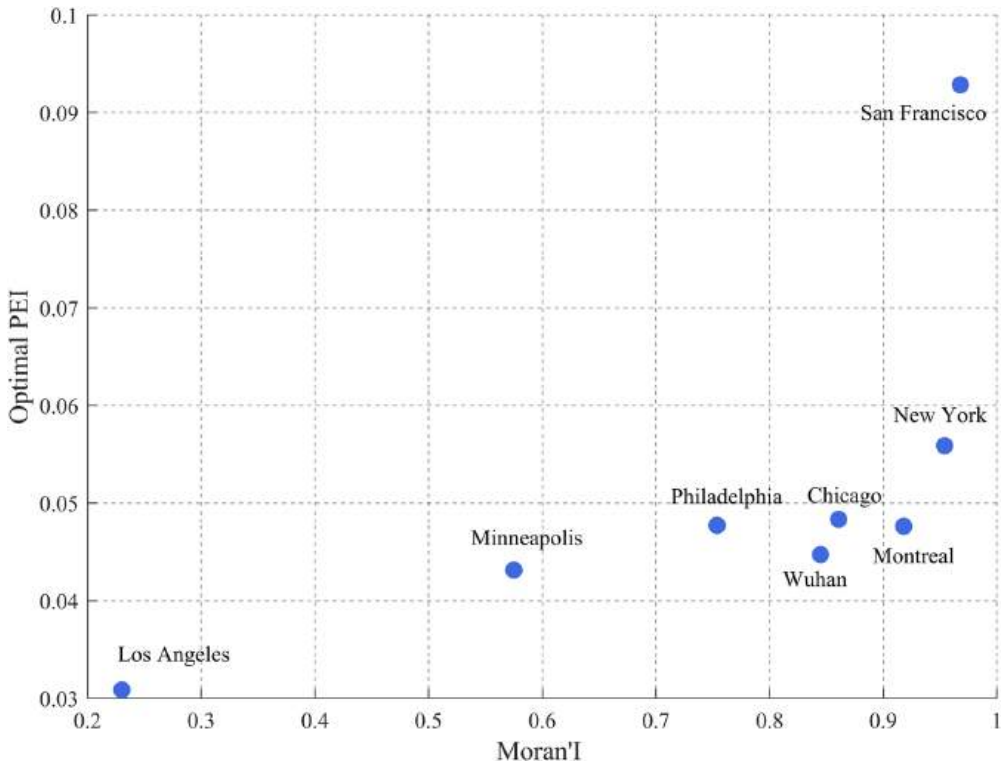

Figure 7. Moran's I and optimal PEI for various cities.

the model. In contrast, cities such as Philadelphia and Minneapolis exhibit broader optimal prediction multiplier ranges with multiple discrete local optimal intervals, demonstrating that the model is more sensitive to parameter fluctuations and exhibits some instability in prediction performance.

In summary, crime spatial clustering not only influences the model's ability to capture effective spatial patterns but also directly determines the optimization space for prediction parameters and the model's stability. In cities with strong spatial clustering, deep learning models can establish more stable and accurate spatio-temporal mapping relationships, thereby achieving better performance at lower threshold settings. Conversely, in cities with weaker clustering, models require higher coverage multiples and more refined parameter tuning strategies to address the uncertainty of hotspot distributions. Therefore, spatial autocorrelation, as an important indicator of the complexity of criminal geographic structures, should be taken into account as a key factor in cross-city model transfer and parameter selection.

3.3.3. Spatial structure types of crime and PEI

To investigate the impact of urban crime spatial structure on predictive performance, this study employed the kernel density analysis tool to assess the spatial distribution of crime incidents in eight cities. Kernel density analysis is employed to calculate the density of spatial point or line features within their surrounding neighbourhoods and to simulate continuous density distributions. By calculating the kernel density values for each grid cell in the image, the distribution characteristics of spatial features can be quantified. This method can reveal

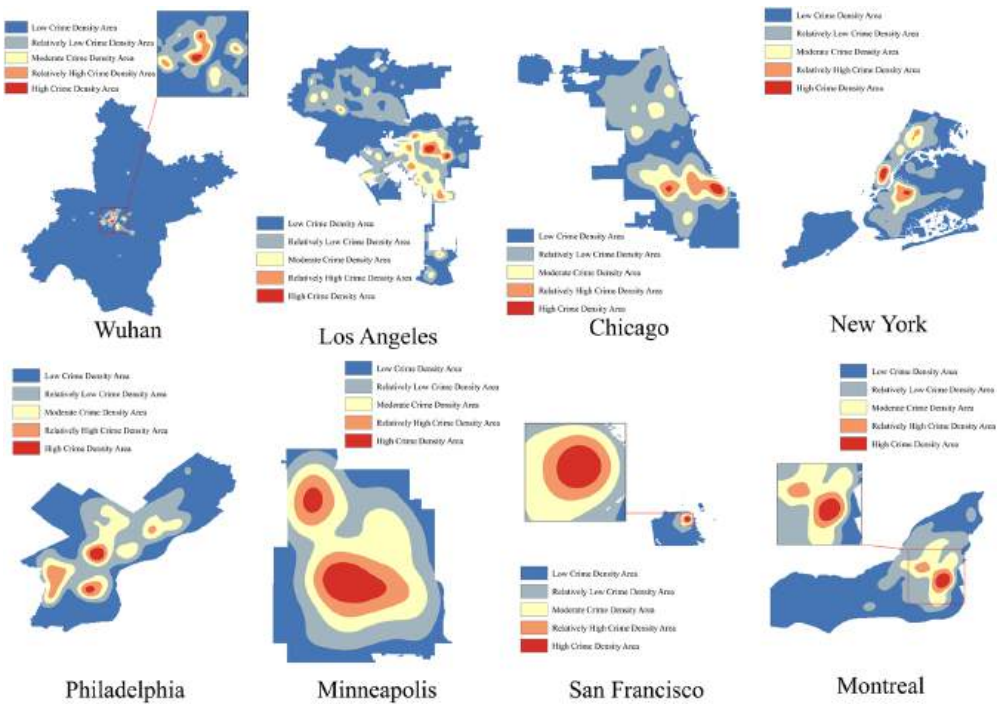

Figure 8. Kernel density analysis results.

the clustering patterns between crime hotspots and their surrounding environments, thereby characterizing the spatial distribution of crime incidents within the study area.

Kernel density analysis was applied to calculate the density of crime hotspots, and the clustering characteristics of crime occurrences were further analyzed based on the spatial distribution of crime hotspots within the study area. A higher kernel density estimate indicates a higher spatial concentration of crime hotspots in the region. The specific calculation method for kernel density analysis is as follows:

$$f(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{31}$$

where $f(x)$ is the kernel density value of the distribution of crime hotspots, n is the number of crime hotspots, h is the bandwidth (set to 2 kilometers in this study), $x - x_i$ is the distance between the crime hotspot x and the sample point x_i, and K is the kernel function. In this study, the Gaussian kernel function is employed, as shown in Equation (32).

$$K\left(\frac{x - x_i}{h}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_i)^2}{2h^2}}. \tag{32}$$

Figure 8 shows the kernel density analysis results for crime occurrence in each city. Based on the kernel density values, the crime density areas of each city are divided into five levels, from low to high: low, relatively low, moderate, relatively high, and high crime density areas.

When analyzing crime density zones in different cities, it was observed that some cities have only one primary high-crime density zone (e.g., San Francisco and Montréal), with

lower crime incidence densities in other areas, exhibiting a typical 'single-centre' spatial structure, whereas other cities have multiple relatively independent high-crime density zones (e.g., New York and Chicago), exhibiting a 'multi-centre' spatial distribution pattern. Based on this structural difference, this paper categorizes cities into 'single-centre cities' and 'multi-centre cities' and examines their impact on predictive performance.

To overcome the subjectivity inherent in relying solely on visual inspection of kernel density estimation (KDE) maps, this study introduces a peak dominance index (PDI) to enable a quantitative distinction between 'single-centre cities' and 'multi-centre cities', while retaining KDE as the fundamental tool for spatial representation. The construction of the PDI is conceptually grounded in the notion of urban primacy proposed by Jefferson (1939) and the quantitative frameworks for polycentric urban structures developed by McMillen (2001). Analogous to primacy indices that assess the dominance of the largest city within an urban system, the PDI is designed to measure the extent to which the dominant crime hotspot on a KDE surface governs the overall distribution of significant hotspots. If a city's crime spatial structure is controlled by a single dominant hotspot, the corresponding KDE surface should exhibit one principal peak whose intensity markedly exceeds that of all secondary centres. Conversely, if multiple hotspots of comparable intensity coexist and are spatially separated, the city should be characterized as having a 'multi-centre' structure.

Specifically, local maxima are first identified on the KDE surface of each city and treated as potential crime hotspot peaks. Because KDE surfaces in continuous space inevitably contain secondary minor peaks induced by local fluctuations or boundary effects, indiscriminately incorporating all local maxima would tend to overestimate the number of centres and dilute the structural significance of the dominant peak. To address this issue, a significant-peak filtering mechanism is introduced, whereby only peaks whose density exceeds a fixed proportion of the global maximum are retained as candidate centres. Let f_k denote the KDE intensity of the kth peak and $f_{\max}$ represent the global maximum peak intensity. The set of significant peaks is then defined by the condition $f_k \geq 0.4 \times f_{\max}$. Adopting a relative threshold rather than an absolute density cutoff eliminates scale differences in KDE values across cities and ensures comparability, while effectively removing non-structural noise peaks that are substantially weaker than the dominant hotspot. This filtering strategy is intended to distinguish structurally meaningful high-incidence areas from random background fluctuations. By imposing a stringent intensity criterion, only secondary centres associated with substantive public safety risks are considered, thereby preventing the misclassification of urban crime structures caused by conflating minor neighbourhood-level incidents with genuine multi centre patterns.

Based on the resulting set of significant peaks, the peak dominance index is defined as the ratio of the global maximum peak intensity to the sum of intensities across all significant peaks:

$$\mathrm{PDI} = \frac{f_{\max}}{\sum_{k=1}^{K} f_k}, \tag{33}$$

where K denotes the number of significant peaks. The PDI takes values in the interval $(0, 1]$. Values approaching 1 indicate that the dominant peak accounts for a large proportion of the total intensity of significant hotspots, corresponding to a single centre spatial structure. In contrast, lower PDI values imply that multiple hotspots contribute comparably to the overall intensity, reflecting a multi centre configuration. For interpretative purposes, this study adopts PDI ≥ 0.6 as the criterion for identifying single centre structures. This threshold is

Table 11. Peak dominance index (PDI) and monocentric classification results.

City	P_{max} (Global Max)	$\sum P_i$ (Valid Sum)	PDI	Single-centre
Wuhan	1,786,266.9	4,171,184.0	0.4282	×
Los Angeles	571,117.3	6,343,553.0	0.0900	×
Chicago	837,409.1	4,181,638.5	0.2003	×
New York	1,007,884.6	3,008,064.0	0.3351	×
Philadelphia	691,435.5	3,688,080.2	0.1875	×
Minneapolis	440,343.0	1,056,819.5	0.4167	×
San Francisco	1,516,817.0	1,516,817.0	1.0000	✓
Montréal	1,199,789.4	1,698,599.6	0.7063	✓

Table 12. Center conditions and PEI for various cities.

City	Single-centre?	Optimal prediction multiples k^*	Optimal PEI
Wuhan	×	2.3	0.044734
Los Angeles	×	3.4	0.030881
Chicago	×	2.7	0.048342
New York	×	2.7	0.055891
Philadelphia	×	3.8	0.047732
Minneapolis	×	5	0.043149
San Francisco	✓	1.6	0.092853
Montréal	✓	1.3	0.047637

motivated by the logic of structural dominance: when a single peak contributes more than 60% of the total intensity of significant hotspots, it is considered to exert overwhelming dominance and the city is classified as a single centre; otherwise, the city is categorized as a multi centre.

Table 11 reports the global maximum peak intensity, the total intensity of significant peaks, and the corresponding PDI values for each city. The results show that San Francisco and Montréal exhibit substantially higher PDI values than the other cities, with only one dominant hotspot present in their significant peak sets. By contrast, cities such as New York, Chicago, and Los Angeles display markedly lower PDI values, indicating that their crime intensity distributions are jointly shaped by multiple hotspots of comparable weight and thus conform to a typical polycentric spatial structure. These quantitative findings are consistent with the visual patterns observed in the KDE maps, while avoiding the uncertainty associated with purely subjective visual classification.

As shown in Table 12, the optimal prediction multiples for multi-centre cities tend to be higher than those for single-centre cities (e.g., 2.7 times for New York, 3.8 times for Philadelphia, and as high as 5 times for Minneapolis), whereas those for single-centre cities like Montréal and San Francisco are only 1.3 and 1.6, respectively. This indicates that in multi-centre cities, crime hotspots are more dispersed, necessitating models to expand their coverage range to achieve better predictive performance; in contrast, crime hotspots in single-centre cities are more concentrated, leading to lower requirements for prediction multiples.

To test whether a significant difference exists in the optimal prediction multiple between multi-centre cities and single-centre cities, the nonparametric Mann-Whitney U test was employed. The test results indicated that the U value for this variable was 12.0, with a P value of 0.065 (Table 13). Based on the median values, the optimal prediction multiple for multi-centre cities was 3.05, which was notably higher than the 1.45 for single-centre cities. Although the P value did not reach the traditional 5% significance level, it remained below 0.1. At the 10% significance level, the difference in optimal prediction multiples between the

Table 13. Mann-Whitney *U* test results for optimal prediction multiples between different types of cities.

Variable	*U* value	*P* value	Median in multi-centre cities	Median in single-centre cities
Optimal Prediction Multiple	12	0.065133	3.05	1.45

Table 14. Spearman correlation and significance test between spatial feature variables and prediction performance indicators.

Spatial feature variable	Prediction performance indicator	Spearman ρ	*p*-value
Ellipse grid area ratio	Optimal prediction multiple	0.838	0.0093**
Ellipse grid area ratio	Optimal PEI	−0.571	0.1390
Moran's I	Optimal prediction multiple	−0.743	0.0349*
Moran's I	Optimal PEI	0.881	0.0039**
Single-centre?	Optimal prediction multiple	−0.76	0.0285*
Single-centre?	Optimal PEI	0.378	0.3559

two types of cities approached statistical significance, suggesting that urban spatial structure may have a certain impact on crime prediction effectiveness. Multi-centre cities are more spatially dispersed, and prediction models require higher coverage multiples to adequately capture all crime hotspots.

In summary, differences in urban central structures of crime significantly influence the performance of crime prediction models. In single-centre cities, the model more effectively identifies concentrated hotspots and operates with higher efficiency, whereas in multi-centre cities, it requires more complex spatial coverage strategies, resulting in greater prediction difficulty and higher resource demands.

This finding provides valuable guidance for urban managers, enabling a better understanding of crime spatial distribution characteristics and facilitating the adjustment of public security management strategies and resource allocation. These insights can enhance social stability and public security while minimizing resource waste. By integrating crime density distribution data, public safety resource allocation can be optimized, and police deployment made more efficient, enabling targeted social interventions and preventive measures in high-risk areas, thereby effectively reducing crime rates.

3.3.4. Correlation analysis of spatial feature variables

To further quantify the relationship between urban spatial structure characteristics and prediction performance, Spearman's rank correlation coefficient (Spearman's ρ) was employed to examine the correlation between each core spatial feature variable and the prediction performance indicators (optimal PEI and optimal prediction multiple). The results are presented in Table 14 and Figure 9.

From the perspective of prediction accuracy, a significant positive correlation was observed between Moran index and optimal PEI ($\rho = 0.881$, $p = 0.0039$), indicating that a stronger spatial clustering of urban crime hotspots corresponds to an improved CNN-LSTM model prediction performance in the city. This result is consistent with the tendency of deep learning models to be more effective in learning structural data.In cities exhibiting highly aggregated spatial structures, crime distributions display pronounced spatial patterns, enabling spatio-temporal sequence-based deep learning algorithms to capture the evolution of crime patterns more effectively. In contrast, the correlation between the elliptical grid area ratio and optimal PEI did not reach statistical significance ($\rho = -0.571, p = 0.1390$), and the association between the single-centre structure and PEI was also weak, suggesting that spatial

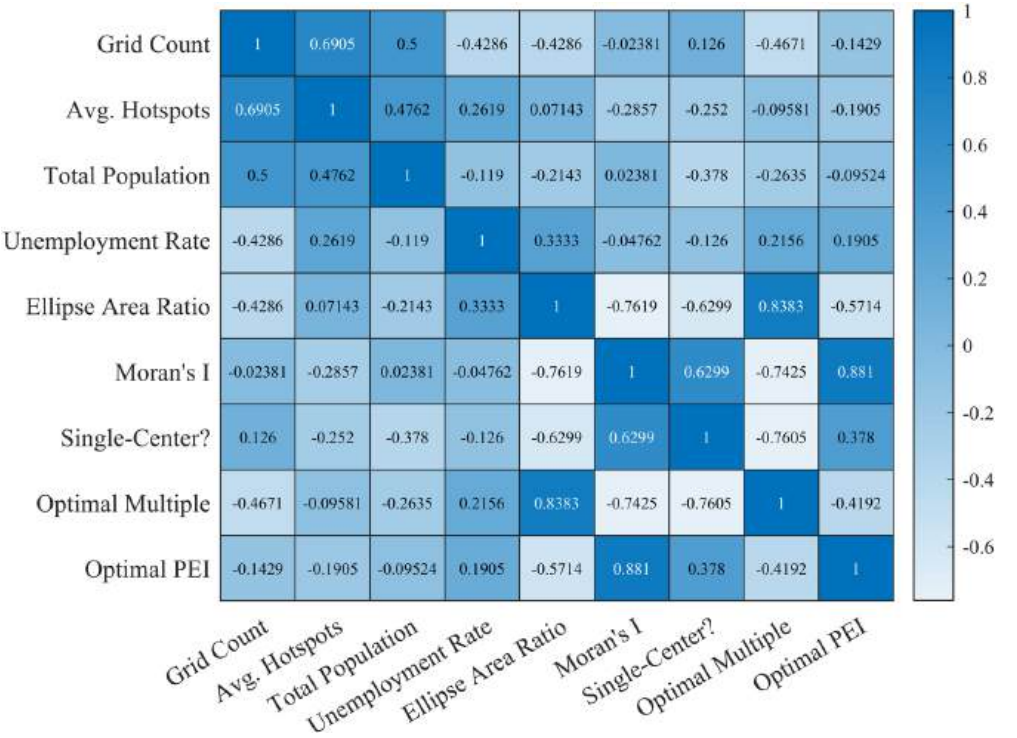

Figure 9. Spearman correlation coefficient heat map between variables.

clustering plays a more critical role in prediction accuracy than the geometric distribution range or the macro-level central structure.

In terms of predictive strategies, correlations between various spatial structure indicators and the optimal predictive multiplier reveal the impact of the urban spatial structure on resource allocation. Moran index exhibits a significant negative correlation with the optimal predictive multiplier ($\rho = -0.743$, $p = 0.0349$), indicating that in cities with strong spatial aggregation, high predictive coverage can be achieved using fewer predictive regions. Similarly, a significant negative correlation was observed between the monocentricity of the city and the optimal prediction multiplier ($\rho = -0.760, p = 0.0285$), further supporting the positive effect of structural concentration on prediction resource efficiency. Additionally, the significant positive correlation between the elliptical grid area ratio and the optimal prediction multiplier ($\rho = 0.838, p = 0.0093$) indicates that, when the crime spatial distribution is more dispersed and geometrically broader, the model must expand the prediction range to maintain high coverage. These findings suggest that the urban spatial form not only influences model prediction accuracy but also plays a key role in the efficiency of prediction resource utilization.

The interrelationships among spatial structural variables also warrant attention. Figure 9 illustrates a moderate negative correlation between Moran index and the elliptical grid area ratio ($\rho = -0.7619$), suggesting that stronger spatial clustering of hotspots leads to a more concentrated spatial distribution, whereas lower clustering results in a more dispersed pattern. The positive correlation between monocentricity and Moran index ($\rho = 0.6299$)

Table 15. Background characteristics of cities.[2]

City	Grid count	Average number of crime hotspots	Total population	Unemployment rate
Wuhan	20,770	35.7368746	10,607,700	3.08%
Los Angeles	3648	38.4734592	3,971,896	7.20%
Chicago	1505	52.1480687	2,720,556	9.50%
New York	276	12.2329114	8,550,405	7.40%
Philadelphia	868	20.8613620	1,567,442	10.90%
Minneapolis	198	10.2133221	410,935	5.70%
San Francisco	2607	13.5133452	864,816	5.10%
Montréal	1320	15.7978548	1,704,694	9.30%

further indicates that monocentric cities generally exhibit higher levels of hotspot clustering. These structural relationships provide insight into urban spatial mechanisms and offer a theoretical basis for developing classification-based prediction strategies tailored to different urban forms.

Based on the above analysis, four commonly used urban background characteristics – namely, the total number of grids, the average number of crime hotspots, the total population, and the unemployment rate (values reported in Table 15) – were further incorporated to assess their impact on predictive performance and spatial structure. However, the correlation analysis results (Figure 9) indicated that the associations between these background variables and spatial feature variables were generally weak. For instance, the correlation coefficient between the total number of grids and Moran index was only −0.024; the correlation between total population and the elliptical area ratio was −0.214; and the correlation between the unemployment rate and Moran index was as low as −0.048, none of which reaches statistical significance. In addition, the associations between these background variables and the optimal prediction multiple as well as the optimal PEI remain weak. Specifically, the correlation between the total number of grids and the optimal prediction multiple was −0.467, while that between the total population and the optimal PEI was only −0.095, indicating an almost negligible linear relationship. These results suggest that macro-level urban attributes, such as population size, employment conditions, and grid division scale, have a substantially smaller impact on crime prediction performance than spatial structural factors. In other words, a city's population size or socioeconomic conditions do not directly determine the spatial clustering of crime or the effectiveness of crime prediction. This finding further highlights the central role of spatial structural characteristics in crime prediction, indicating that differences in spatial distribution patterns constitute the key determinants of model performance and resource utilization efficiency.

Overall, the correlation analysis corroborates the empirical findings presented above regarding the influence of spatial structure on predictive performance. In cities characterized by stronger crime spatial clustering and more concentrated spatial structures, the CNN–LSTM model not only achieves higher predictive accuracy but also requires fewer predictive resources, thereby exhibiting superior cost efficiency. Moreover, significant associations among different spatial feature variables are observed, reflecting the systematic influence of urban structural factors on crime prediction tasks. By contrast, the effects of background variables, such as population size and unemployment rate, are relatively limited, further underscoring the practical value of spatial feature analysis in urban safety governance and the formulation of crime prediction strategies.

[2] All urban background feature data used in this paper are from 2015. The total population data for Wuhan are obtained from the 'Wuhan City 2016 Statistical Yearbook', while the unemployment rate data are sourced from the 'Statistical Bulletin on the

4. Summary

This paper investigates burglary prediction across multiple cities in different countries, with a particular focus on prediction efficiency rather than accuracy alone. Using a unified CNN–LSTM framework under a consistent spatiotemporal scale, we examine burglary prediction in eight representative cities in China, the United States, and Canada, and evaluate performance using precision rate, hit rate, prediction accuracy index (PAI), and prediction effectiveness index (PEI). After validating the robustness of PEI as an efficiency-oriented evaluation metric, the study integrates spatial structural measures – including standard deviation ellipses, Moran's I, kernel density estimation, and centre dominance analysis – to explore how urban spatial organization shapes predictive performance and resource utilization.

Overall, the empirical results demonstrate that PEI provides a stable and controllable criterion for balancing prediction accuracy and coverage, and that each city exhibits an optimal prediction multiplier that maximizes prediction efficiency. More importantly, the findings consistently indicate that urban spatial structure plays an important role in determining both predictive performance and resource efficiency. Cities characterized by stronger spatial clustering and more concentrated crime distributions tend to achieve higher predictive effectiveness with lower prediction multipliers, whereas cities with more dispersed or polycentric crime patterns require higher multipliers and broader coverage to attain comparable performance. These patterns indicate that prediction efficiency is strongly associated with spatial organization rather than by model specification alone.

Beyond the specific empirical results, this study contributes a transferable analytical framework that systematically links prediction efficiency, spatial structure, and resource allocation strategies. The proposed framework is designed to be adaptable across different urban contexts, as it relies on generic inputs such as event locations and spatial aggregation units. Moreover, although a CNN–LSTM model is adopted for empirical demonstration, the framework itself is model-agnostic and can be readily extended to alternative architectures, including graph-based models or spatiotemporal transformer frameworks. From an applied perspective, the results suggest that incorporating spatial structural diagnostics into predictive modelling can support more informed and cost-effective policing strategies, by aligning prediction intensity with the underlying spatial organization of crime.

Future research can be further advanced in the following directions.

(1) Expansion of data dimensions: The integration of multi-source data spanning social, economic, and cultural domains may enable a more comprehensive characterization of the multifaceted factors influencing crime occurrence, thereby improving the interpretability and accuracy of predictive models.
(2) Incorporation of temporal dynamics in spatial analysis: The current study focuses on relatively stable spatial structures derived from annual aggregated crime data. Future research may introduce dynamic spatial indicators or time-varying hotspot structures (e.g., monthly or quarterly spatial features) to further investigate how temporal variations in crime distribution influence prediction performance.

National Economic and Social Development of Wuhan in 2015 and during the 12th Five-Year Plan period'. The relevant data for Los Angeles, Chicago, New York, Philadelphia, Minneapolis, and San Francisco are respectively sourced from the U.S. Census Bureau and the Bureau of Labour Statistics. The total population and unemployment rate data for Montréal are sourced from Statistics Canada.

(3) Expansion of sample scope: Future studies may include a wider range of cities with different socioeconomic conditions and urban development levels to further evaluate the generalizability of the proposed framework.
(4) Diversification of algorithmic frameworks: A systematic comparison of the performance of various machine learning and deep learning models in crime prediction may help to assess the robustness and applicability of the conclusions.
(5) Extension of application scenarios: The integrated application of the proposed methods in real-time prediction and dynamic police force deployment may be explored to evaluate their practical effectiveness in rapid response and resource optimization.

In summary, rather than treating crime prediction as a purely algorithmic optimization problem, this study emphasizes the integration of spatial structure into efficiency-oriented prediction and decision-making. By explicitly linking spatial concentration patterns with optimal prediction strategies, the proposed framework provides a systematic analytical approach that can be applied to examine how urban spatial organization constrains predictive performance and resource efficiency. Beyond its methodological contributions to crime prediction research, this framework also offers practical guidance for intelligent, refined, and forward-looking public safety management, thereby demonstrating considerable social relevance and broad application potential across diverse urban contexts.

Acknowledgements

We are grateful to the referees for their insightful comments that have led to a substantial improvement to an earlier version of the paper.

Author contributions

CRediT: **Rui Wang:** Data curation, Formal analysis, Writing – review & editing; **Yaofeng Zhang:** Conceptualization, Methodology, Project administration, Supervision; **Jinling Yao:** Data curation, Validation, Visualization, Writing – original draft

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Key Project of the National Social Science Fund of China [grant number 23ATJ005].

References

Alves, L. G. A., Ribeiro, H. V., & Rodrigues, F. A. (2018). Crime prediction through urban metrics and statistical learning. *Physica A: Statistical Mechanics and Its Applications*, *505*, 435–443. https://doi.org/10.1016/j.physa.2018.03.084

Chainey, S., & Ratcliffe, J. (2005). Identifying crime hotspots. In *Gis and crime mapping* (pp. 145–182). John Wiley & Sons. https://doi.org/10.1002/9781118685181.ch6

Chainey, S., Tompson, L., & Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, *21*(1), 4–28. https://doi.org/10.1057/palgrave.sj.8350066

Esquivel, N., Nicolis, O., Peralta, B., & Mateuet, J. (2020). Spatio-temporal prediction of Baltimore crime events using CLSTM neural networks. *IEEE Access*, *8*,209101–209112. https://doi.org/10.1109/Access.6287639

Gu, H., Chen, P., & Li, H. (2021). Overview and prospect for spatial-temporal prediction of crime (in Chinese). *Journal of Geo-information Science*, *23*(1), 43-–57. https://doi.org/10.12082/dqxxkx.2021.200247

He, R., Lu, Y., Jiang, C., Deng, Y., Li, X., & Shi, D. (2023). Progress in research and practice of spatial-temporal crime prediction over the past decade (in Chinese). *Journal of Geo-information Science*, *25*(4), 866-–882. https://doi.org/10.12082/dqxxkx.2023.230299

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu, T., Zhu, X., Duan, L., & Guo, W. (2018). Urban crime prediction based on spatio-temporal Bayesian model. *PloS One*, *13*(10), e0206215. https://doi.org/10.1371/journal.pone.0206215

Jefferson, M. (1939). The law of the primate city. *Geographical Review*, *29*(2), 226–232. https://doi.org/10.2307/209944

Jin, G., Wang, Q., Zhu, C., Feng, Y., Huang, J., & Zhou, J. (2020). Addressing crime situation forecasting task with temporal graph convolutional neural network approach. In *2020 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)* (pp. 474–478). IEEE. https://doi.org/10.1109/ICMTMA50254.2020.00108

Kianmehr, K., & Alhajj, R. (2006). Crime hot-spots prediction using support vector machine. In *IEEE International Conference on Computer Systems and Applications* (pp. 952–959). IEEE Computer Society. https://doi.org/10.1109/AICCSA.2006.205203

Kianmehr, K., & Alhajj, R. (2008). Effectiveness of support vector machine for crime hot-spots prediction. *Applied Artificial Intelligence*, *22*(5), 433–458. https://doi.org/10.1080/08839510802028405

Law, J., Quick, M., & Chan, P. (2014). Bayesian spatio-temporal modeling for analysing local patterns of crime over time at the small-area level. *Journal of Quantitative Criminology*, *30*(1), 57–78. https://doi.org/10.1007/s10940-013-9194-1

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (2002). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Lefever, D. W. (1926). Measuring geographic concentration by means of the standard deviational ellipse. *American Journal of Sociology*, *32*(1), 88–94. https://doi.org/10.1086/214027

Li, M., Du, W., Wen, S., Li, Q., Zhang, T., & Zhong, W. (2025). Crime forecasting: A spatio-temporal analysis with deep learning models. *Journal of Computational Methods in Sciences and Engineering*, *25*(5), 4090–4099. https://doi.org/10.1177/14727978251337993

Liu, L., Liu, W., Liao, W., Yu, H., Jiang, C., Lin, R., & Zhang, Z. (2018). Comparison of random forest algorithm and space-time kernel density mapping for crime hotspot prediction (in Chinese). *Progress in Geography*, *37*(6), 761-–771. https://doi.org/10.18306/dlkxjz.2018.06.003

Maheshwari, A., Camargo, M. E., Sathiyanarayanan, M., & Filho, W. P. (2025). Urban vs. rural crime dynamics: A spatio-temporal and machine learning approach for comparative analysis. In *2025 International Conference on Frontier Technologies and Solutions (ICFTS)* (pp. 1–10). IEEE. https://doi.org/10.1109/ICFTS62006.2025.11031983

Matereke, T., Nyirenda, C. N., & Ghaziasgar, M. (2021). A comparative evaluation of spatio-temporal deep learning techniques for crime prediction. In *2021 IEEE Africon* (pp. 1–6). IEEE. https://doi.org/10.1109/AFRICON51333.2021.9570858

McMillen, D. P. (2001). Nonparametric employment subcenter identification. *Journal of Urban Economics*, *50*(3), 448–473. https://doi.org/10.1006/juec.2001.2228

Mei, Y., & Li, F. (2019). Predictability comparison of three kinds of robbery crime events using LSTM. In *Proceedings of the 2019 2nd International Conference on Data Storage and Data Engineering* (pp. 22–26). Association for Computing Machinery. https://doi.org/10.1145/3354153.3354162

Michael, T., Ross, H., & Janet, C. (2003). Infectious burglaries. A test of the near repeat hypothesis. *The British Journal of Criminology*, *43*(3), 615–633. https://doi.org/10.1093/bjc/43.3.615

Mohler, G. O., Short, M. B., & Brantingham, P. J. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, *106*(493), 100–108. https://doi.org/10.1198/jasa.2011.ap09546

Mohler, G. O., Short, M. B., & Brantingham, P. J. (2017). The concentration-dynamics tradeoff in crime hot spotting. In *Unraveling the Crime-Place Connection* (pp. 19–39). Routledge.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17–23. https://doi.org/10.2307/2332142

Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, *33*(3), 1065–1076. https://doi.org/10.1214/aoms/1177704472

Pimple, S., Mansoor, H. D., & Karthikeyan, N. (2023). Enhanced public safety through LSTM-based spatio-temporal crime analytics. In *2023 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICERCS57948.2023.10434246

Polvi, N., Looman, T., Humphries, C., & Pease, K. (1991). The time course of repeat burglary victimization. *The British Journal of Criminology*, *31*(4), 411–414. https://doi.org/10.1093/oxfordjournals.bjc.a048138

Ratcliffe, J. (2009). Crime mapping: Spatial and temporal challenges. In *Handbook of quantitative criminology* (pp. 5–24). Springer.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, *27*(3), 832–837. https://doi.org/10.1214/aoms/1177728190

Roshankar, R., & Keyvanpour, M. R. (2023). Spatio-temporal graph neural networks for accurate crime prediction. In *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)* (pp. 168–173). IEEE. https://doi.org/10.1109/ICCKE60553.2023.10326223

Rummens, A., & Hardyns, W. (2020). Comparison of near-repeat, machine learning and risk terrain modeling for making spatiotemporal predictions of crime. *Applied Spatial Analysis and Policy*, *13*(4), 1035–1053. https://doi.org/10.1007/s12061-020-09339-2

Rummens, A., Hardyns, W., & Pauwels, L. (2017). The use of predictive analysis in spatiotemporal crime forecasting: Building and testing a model in an urban context. *Applied Geography*, *86*, 255–261. https://doi.org/10.1016/j.apgeog.2017.06.011

Shen, H., Zhang, H., Zhang, Y., Zhang, Z., Zhu, Y., & Cai, L. (2019). Prediction of burglary crime based on LSTM (in Chinese). *Journal of Statistics and Information*, *34*(11), 107–115. https://doi.org/10.3969/j.issn.1007-3116.2019.11.014

Short, M. B., Mohler, G. O., Brantingham, P. J., & Tita, G. E. (2014). Gang rivalry dynamics via coupled point process networks. *Discrete and Continuous Dynamical Systems – Series B*, *19*(5), 1459–1477. https://doi.org/10.3934/dcdsb.2014.19.1459

Wang, S., Lin, Y., Jia, Y., Sun, J., & Yang, Z. (2024). Unveiling the multi-dimensional spatio-temporal fusion transformer (MDSTFT): A revolutionary deep learning framework for enhanced multivariate time series forecasting. *IEEE Access*, *12*,115895–115904. https://doi.org/10.1109/ACCESS.2024.3444788

Zhang, Y., Yao, J., Li, L., & Geng, Z. (2023). Construction and application of crime prediction effectiveness index (in Chinese). *Mathematics in Practice and Theory*, *53*(12), 163–175.

Zhao, D., Du, P., Liu, T., & Ling, Z. (2023). Spatio-temporal distribution prediction model of urban theft by fusing graph autoencoder and GRU (in Chinese). *Journal of Geo-Information Science*, *25*(7), 1448–1463. https://doi.org/10.12082/dqxxkx.2023.220584