Journal of Environmental Management 317 (2022) 115438

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Assessing myocardial infarction severity from the urban environment perspective in Wuhan, China

Yao Yao [a], Hanyu Yin [b], Changwu Xu [c,d,e], Dongsheng Chen [f,*], Ledi Shao [a], Qingfeng Guan [a], Ruoyu Wang [g]

[a] School of Geography and Information Engineering, China University of Geosciences, Wuhan, 430078, Hubei province, PR China
[b] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430072, Hubei province, PR China
[c] Department of Cardiology, Renmin Hospital of Wuhan University, Wuhan, 430060, PR China
[d] Cardiovascular Research Institute, Wuhan University, Wuhan, 430060, PR China
[e] Hubei Key Laboratory of Cardiology, Wuhan, 430060, PR China
[f] China Regional Coordinated Development and Rural Construction Institute, Sun Yat-sen University, Guangzhou, 510000, Guangdong Province, PR China
[g] UKCRC Centre of Excellence for Public Health/Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom

ARTICLE INFO

Keywords:
Myocardial infarction
Interpretable machine learning
Spatio-temporal factors
GIS spatial Analysis
Urban environment

ABSTRACT

Health inequalities are globally widespread due to the regional socioeconomic inequalities. Myocardial infarction (MI) is a leading health problem causing deaths worldwide. Yet medical services for it are often inequitably distributed by region. Moreover, studies concerning MI's potential spatial risk factors generally suffer from difficulties in focusing on too few factors, inappropriate models, and coarse spatial grain of data. To address these issues, this paper integrates registered 1098 MI cases and urban multi-source spatio-temporal big data, and spatially analyses the risk factors for MI severity by applying an advanced interpretable model, the random forest algorithm (RFA)-based SHapley Additive exPlanations (SHAP) model. In addition, a community-scale model between spatio-temporal risk factors and MI cases is constructed to predict the MI severity of all communities in Wuhan, China. The results suggest that those risk factors (i.e., age of patients, medical quality, temperature changes, air pollution and urban habitat) affect the MI severity at the community scale. We found that Wuhan residents in the downtown area are at risk for high MI severity, and the surrounding suburb areas show a donut-shape pattern of risk for medium-to-high MI severity. These patterns draw our attention to the impact of spatial environmental risk factors on MI severity. Thus, this paper provides three recommendations for urban planning to reduce the risk and mortality from severe MI in the aspect of policy implication.

1. Introduction

Improving people's average health and reducing health inequalities are humanly desirable. Currently, MI is the leading health problem causing cases of death worldwide (White and Chew, 2008). Acute MI is defined as myocardial necrosis caused by acute, sustained ischemia and hypoxia in the coronary arteries (W.H.O. 1988). Nearly 16 million people worldwide suffered from MI in 2015 (Vos et al. 2016). And previous studies found that regional differences often exist in the incidence of the disease. It is most common in developed countries with good welfare, where the mortality rate for ST-segment elevation MI is around 10% (Members et al. 2012). However, it has also been shown that the risk of death from cardiovascular disease (CVD) tends to be lower in areas of better socio-economic status, i.e., in areas with greener housing and closer to healthcare services (Chen et al. 2020; Widimský et al. 2003). Thus, it is imperative to examine the potential risk factors behind MI severity in order to protect residents' health and reduce mortality.

Suffering from severe MI may be influenced by numerous risk factors (Nawrot et al. 2011; Smyth et al. 2016). Previous studies suggest that risk factors for MI severity may focus on three main dimensions: the physical environment (e.g., temperature changes, air pollution, traffic exposure) (Chen et al. 2019; Miller et al. 2007; Nieuwenhuijsen, 2018; Raziani and Raziani, 2021; Vienneau et al. 2019), the social status (e.g.,

* Corresponding author.
E-mail addresses: yaoy@cug.edu.cn (Y. Yao), hanyu_yin@whu.edu.c (H. Yin), xuchangwu@whu.edu.cn (C. Xu), dontsingchen@whu.edu.cn (D. Chen), sldmimose@yeah.net (L. Shao), guanqf@cug.edu.cn (Q. Guan), r.wang-54@sms.ed.ac.uk (R. Wang).

https://doi.org/10.1016/j.jenvman.2022.115438
Received 9 November 2021; Received in revised form 8 May 2022; Accepted 27 May 2022
Available online 30 May 2022

residents' neighbourhoods, social networks) (Gerber et al. 2010; Wang et al. 2019) and patients' own body conditions (e.g., smoking and drinking, vigorous exercise, obesity) (Mostofsky et al. 2015; Smyth et al. 2016; Xu et al. 2018). However, limited by the lack of spatial data, some studies have not considered the spatial agglomeration effect on the increased or diminished severity for MI. With the increasing sophistication of technologies related to Geographic Information System (GIS), some studies have explored the use of GIS-based spatial statistical analysis methods for the spatial agglomeration effect to further analyse risk factors for MI severity (Franch-Pardo et al. 2020; Rushton, 2003; Timonin et al. 2018). Thus, the studies are able to help assess disease severity for public health planning to reduce health inequalities.

However, current GIS-based studies of MI generally suffer from difficulties in focusing on too few factors, poor levels of data mining, and coarse-grained spatial data. Most studies have analysed several spatial risk factors from poor data sources, e.g., temperature and air quality map (Akbarzadeh et al. 2018; Liu et al. 2018; Rowland et al. 2020; Yang et al. 2017). Although these studies have provided basic correlations for MI-related public health planning, they lack the ability to consider the gain effect when multiple spatial risk factors act together, thereby failing to comprehensively simulate the ground-truth space. Meanwhile, most studies are also limited by the coarse-grained spatial data and unable to analyse spatial risk factors at microscopic scales (Amsalu et al. 2019; Tonne et al. 2007). Thus, they cannot effectively address the problem of targeted distribution of medical facilities. To deeper investigate the potential spatial risk factors for MI, it is necessary to comprehensively model the complex residential space at microscopic scales.

To address these issues, this paper quantifies the spatial risk factors and their spatial agglomeration effect on MI severity at community-scale. The RFA-based SHAP interpretable model and multi-source spatio-temporal datasets are applied, coupled with community-scale MI patient attendance data, to deeply explore the spatial risk factors. The model considers the positive and negative impact on MI severity to deeply reveal the relationship. At the same time, the severity of MI at the urban community scale is predicted based on GIS spatial analysis. And its spatial distribution pattern is analysed, to achieve the purpose of preventing mortality due to severe MI at the urban micro scale.

2. Literature review

2.1. Spatial risk factors for myocardial infarction

Previous studies have demonstrated that the MI severity shows some spatial patterns. Some studies have explored the specific causes of death from CVDs, e.g., the impact of greenness in residential areas on CVDs (Asri et al. 2020; Yang et al. 2020). In five of the eight relevant literatures, we found a significant association between CVD mortality and residential greenery (Hu et al. 2008; Lachowycz and Jones, 2014; Mitchell and Popham, 2008; Richardson et al. 2010, 2012; Richardson and Mitchell, 2010; Tamosiunas et al. 2014; Villeneuve et al. 2012). However, these studies were limited by the coarse grain of data, and had to analyse the association at coarse spatial scale, which cannot simulate the real space during a CVD attack. Thanks to development of computer vision techniques and the availability of street-view data, recent studies have been able to simulate the complicated space from a human-scale perspective (Helbich et al. 2019).

Spatial risk factors for MI severity can be further summarised as spatial heterogeneity due to socio-economic inequalities. Some found that the effect of SVG on ischemic heart diseases varies significantly by individual demographic and socioeconomic characteristics (Yao et al. 2022). Previous studies presented that the clusters of residents with a risk of low MI severity are more likely to be distributed in the most socio-economically favourable areas (Kihal-Talantikite et al. 2017). The residential space in these areas is characterised by shorter distances to healthcare facilities and greater access to public transport. Spatial heterogeneity due to the socio-economic status may be reflected at more aspects. Thus, spatial risk factors for MI severity need to be further explored.

2.2. Related spatial analysis methods

Many studies in the field of health geography have explored spatial risk factors for MI severity using different spatial analysis methods. A number of studies have analysed the spatial distribution of MI severity considering a single aspect of potential spatial risk (Rowland et al. 2020; Yang et al. 2017). They only analysed the impact of temperature on MI severity, lacking a comprehensive analysis of multi-source spatio-temporal data. Some studies used a time-stratified case-crossover method to statistically analyse the effect of air pollutants such as PM10 and PM2.5 on the risk of suffering MI (Akbarzadeh et al. 2018; Liu et al. 2018). But it's greatly influenced by the setting of time interval and individual exposure differences, thereby suffering from subject selectivity bias and not suitable to consider the impact of long-term living environment (Maclure and Mittleman, 2000).

Some studies explored the spatial agglomeration effect of the potential risk on MI severity at a coarse spatial scale. A previous research used the distance from residential areas to main roads as a proxy for residents' exposure to traffic-related air pollutants (Tonne et al. 2007). However, the impact of traffic exposure is not directly analysed on the MI severity at the aspect of traffic pollutants. Previous studies only analysed the spatial distribution of MI severity at the district and county scale, but not at micro scale (Amsalu et al. 2019). Thus, these results could not effectively address the issue of targeted distribution of medical supplies.

In summary, spatial studies of MI severity generally suffer from difficulties in focusing on too few factors, poor levels of data mining, and lack of fine-grained prediction and mapping.

3. Materials and methodology

3.1. Study area

Wuhan, the capital of Hubei Province, China (Fig. 1 (A) (B)), is selected as the study area for this paper. Wuhan is located in central China, between 29° 58' - 31° 22′ N and 113° 41' - 115° 05′ E. It covers an area of 85.69 million km^2, with a resident population of 11.21 million in 2019 (Statistics, 2020). Wuhan has 13 administrative districts, of which the central Hankou area (Jianghan District, Jiangan District, Hanyang District and Qiaokou District) is the well-developed area with a high density of urban population. Wuchang District and Hongshan District are the areas where universities gather. And Qingshan District is Wuhan's industrial-functioned area. The districts of East-West Lake, Hannan, Caidian, Jiangxia, Huangpi and Xinzhou are suburb areas far away from the central urban. According to statistics, CVDs are the leading causes of death in Wuhan, with about 13 out of every 10,000 people dying from the diseases (Statistics, 2020).

3.2. Data

3.2.1. Myocardial infarction patient attendance data

The data of MI patient cases in Wuhan are collected from the People's Hospital of Hubei Province, containing information on 1098 anonymous MI patient attendances between 2016 and 2019. Table S1 shows the anonymised patient attendance data. It included information on time of admission, current address (latitude, longitude, address), age, gender, number of days in hospital, total cost of hospitalisation, whether or not they died, and the corresponding diagnostic modality for different patients. In this paper, it is hypothesized that the higher the patients' hospital expenditure, the more severe MI that he/she suffers. Here, five levels (0–4) of MI severity are expressed in terms of patients' hospital expenditure with based on the natural discontinuity grading method (Chen et al. 2013). In addition, MI mortality is used in form of a

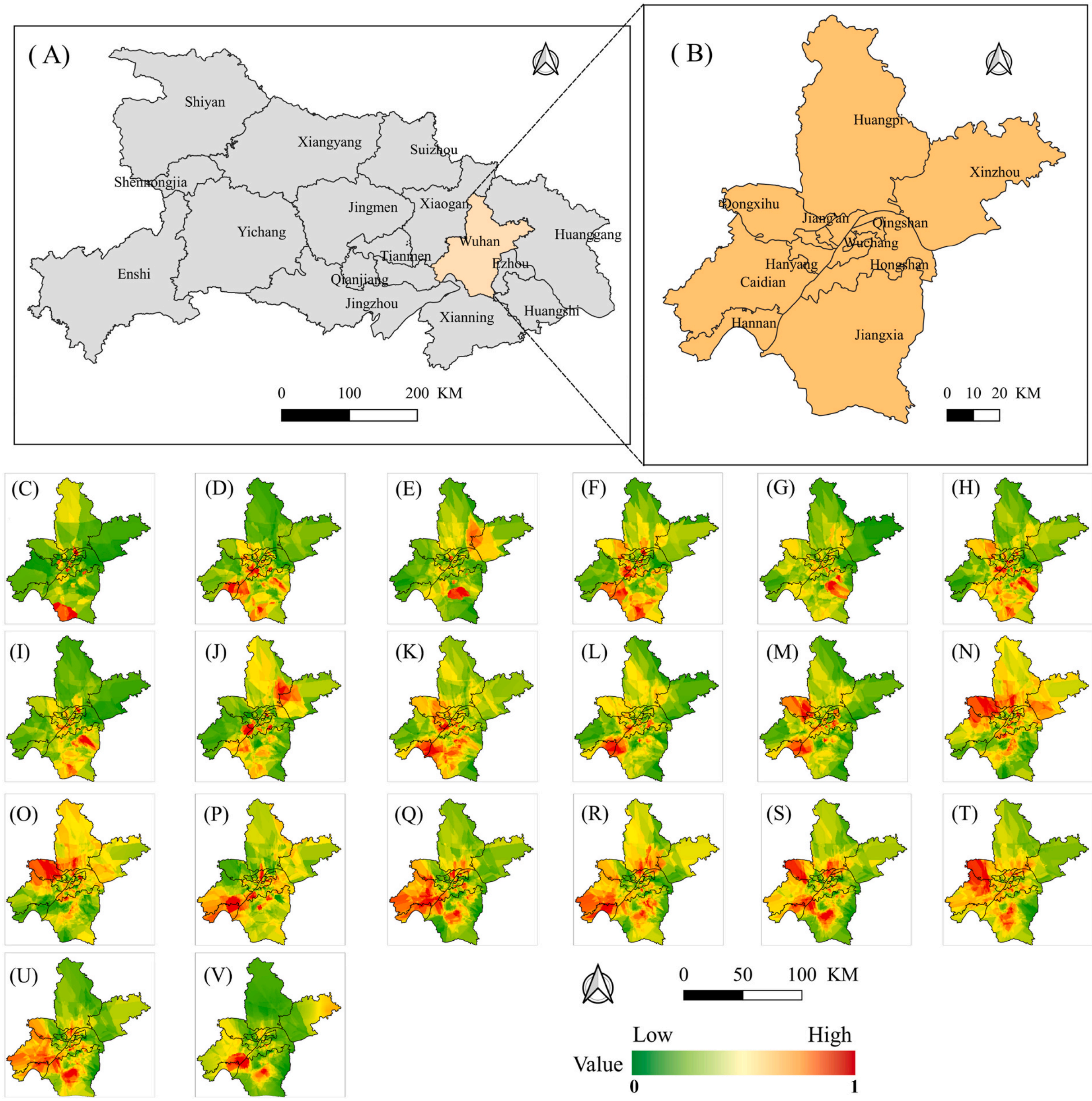

Fig. 1. Location of the study area. (A) Hubei Province, (B) Wuhan City. And the spatio-temporal variables, including (C) famous tourist attractions, (D) scientific, educational and cultural services, (E) car services, (F) public facilities, (G) place name address information, (H) domestic services, (I) medical services, (J) road intersections, (K) night-time population density, (L) regional population density, (M) GDP, (N) maximum temperature, (O) minimum temperature, (P) temperature difference between days, (Q) SO2, (R) NO2, (S) CO, (T) O3, (U) PM10, (V) PM2.5.

dichotomous variable ('0' = survival, '1' = death).

3.2.2. *Multi-source spatio-temporal data*

In this study, Point-of-interest (POI) and OpenStreetMap (OSM: http://openstreetmap.org) road network data are collected. POI data is able to reflect the socio-economic conditions of different regions and help to explore the pattern of urban space (Yao et al. 2017; Yuan et al. 2014). OSM road network data have the advantages of high presentability, rich information, and large data volume, which can well record the key urban traffic data (Goodchild, 2007).

In terms of demographic and economic data, this paper uses Real-time Tencent user density (RTUD) data from 2016 to 2019. RTUD data record the location of Tencent users every hour, of which about 570 million users use WeChat daily, more than 1/3 of China's population. So RTUD data can effectively reflect the range of activities of residents at different times of the day (Yao et al., 2017b).

This paper used meteorological data before and after patient admissions during the period 2016–2019 from the Public Meteorological Service Centre of the China Meteorological Administration (http://www.weather.com.cn/), to represent the climate and weather conditions. This includes temperature data, AQI index and concentrations of various air pollutants (SO2, NO2, CO, O3, PM10, PM2.5) before and after the patients' admission.

For the urban space aspect, Tencent streetview images (https://map.qq.com/jiejing) are used to assess the exposure rates of different categories of features. Based on OSM's road grid sampling points, streetview images from four different angles (0°, 90°, 180°, 270°) were collected. A deep learning model, a fully convolutional neural network (FCN-8s), is used and trained from ADE20K data to identify various categories of feature exposure (Helbich et al. 2019). The model achieved an end-to-end accuracy of 0.81 on the training dataset and 0.67 on the test dataset (Kang and Wang, 2014; Yao et al. 2019). In addition, Street View

Greenness (SVG) is calculated for the landscape (Wang et al. 2020).

Here, descriptive statistics of the above datasets are conducted in Table 1. The average value and standard deviation of these variables are calculated within a 600 m buffer. Moreover, we normalized the data by scaling the range of variable values to 0 (low) - 1 (high). And their spatial distributions are shown in Fig. 1(C)–(V). Areas with high values of the variables are in red and vice versa in green.

3.3. Methodology

The detailed workflow of this paper is shown in Fig. 2. Two steps are included. (1) Data pre-processing. First, the invalid records and incorrect data of the multi-source spatio-temporal data are filtered out. And then, the spatio-temporal risk factors are expressed in terms of the average values of the spatio-temporal variables within a 600 m buffer. (2) Analysis of MI severity. The RFA-SHAP model is trained first by using the multi-source spatio-temporal features and the patient cases data. The performance of the models is evaluated by comparing with conventional statistical regression analysis methods. Next, the prediction of MI severity in Wuhan is mapped and its distribution patterns are analysed.

3.3.1. RFA-based SHAP interpretable model

The RFA has the advantage of incorporating high-dimensional spatio-temporal data and handling thousands of explanatory variables (Breiman, 2001; Lin and Jeon, 2006). In the field of public health,

Table 1
Descriptive statistics for multi-source spatio-temporal data, including the average value, standard deviation and abbreviations of the variables after statistical analysis of the multi-source spatio-temporal data within a 600 m buffer.

Category	Variable	Average value (STD)	Abb
Patient Personal Information	Age (years)	66.56 (±13.208)	
	Male (%)	74.3	
	Female (%)	25.7	
	Recovery (%)	90.6	
	Dead (%)	9.4	
	Disease severity level (0–4)	1.27 (±1.182)	
Multi-source Geographic Data (pcs)	Famous Tourist Sites	7.15 (14.025)	FTS
	Science, Education and Culture Services	70.54 (61.356)	SEC
	Car Service	15.31 (14.914)	CAS
	Public Facilities	11.73 (12.405)	PBF
	Toponymic Address Information	197.68 (186.577)	TAI
	Life Service	229.69 (173.129)	LFS
	Healthcare Service	47.67 (42.891)	HCS
	Road Intersections	131.58 (97.712)	RIS
Demographic Economic Data	Nighttime Population Density (persons/km^2)	8613.47 (3733.519)	NPD
	Regional Population Density (persons/km^2)	7441.64 (4043.206)	RPD
	Neighbourhood GDP (10,000 Yuan/km2)	48247.11 (32589.214)	GDP
Meteorology Data	Maximum Temperature (°C)	23.42 (9.102)	MAT
	Minimum Temperature (°C)	14.45 (9.065)	MIT
	Temperature Difference Between Days (°C)	0.23 (2.480)	TDB
	SO_2 (ug/m^3)	3.79 (2.062)	
	NO_2 (ug/m^3)	23.60 (12.083)	
	CO (mg/m^3)	10.08 (2.933)	
	O_3 (ug/m^3)	22.81 (14.827)	
	PM_{10} (ug/m^3)	60.54 (25.870)	
	$PM_{2.5}$ (ug/m^3)	62.98 (39.957)	
Street View Data (%)	Street-View Green	0.17 (0.063)	SVG
	Street-View Wall	0.03 (0.012)	SVW
	Sky View Factor	0.13 (0.058)	SVF
	Street-View Window	0.00035 (0.0006)	SWI
	Street-View Street	0.18 (0.034)	SVS

Abb = Abbreviation.

previous studies have applied RFA model to analyse the spread risks and drivers of infectious diseases (Yao et al. 2021). While SHAP is an interpretable machine learning method based on game theory (Lundberg and Lee, 2017). This method can calculate the Shapley value corresponding to each variable based on the Monte Carlo sampling method (Rubinstein and Kroese, 2016). Thus, it can quantify the effects of dependent variables on independent variables, thereby bridging the gap of the RFA in terms of variable interpretability.

To analyse the performance of RFA-SHAP in multi-classification tasks like the MI severity analysis, three models (i.e., multiple linear regression, geographically weighted regression (GWR), and multiscale geographically weighted regression (MGWR) (Fotheringham et al. 2017)) are applied for comparison. R^2 (Nakagawa and Schielzeth, 2013) is used for regression accuracy assessment. Also, for classification accuracy assessment, metrics such as precision, recall, and F1 value are used to compare diverse machine learning models (Flach and Kull, 2015). In the MI mortality analysis, given the spatio-temporal autocorrelation in the data, Cox proportional hazards regression model, Cox frailty model, and MGWR model are used for baseline comparisons to analyse the performance of RFA-SHAP in such a binary classification task. R^2 is not a good enough measure for binary classification tasks (Cox and Wermuth, 1992). Thus, AIC (Akaike information criterion) is applied to assess the goodness of fit of the models (Akaike 1974), while C-index is for assessing the prediction accuracy of the models. It should be noted that in the MI mortality analysis of the Cox models, the C-index metric is equivalent to the AUC (area under curve) (Brentnall et al. 2015, Harrell 2015).

Considering the five levels of MI severity in the MI patient attendance data, Equation (1) was used to calculate the model accuracy (MA) for the RFA model.

$$MA = \frac{\sum_{i=1}^{n} \frac{1}{|y_{i-pred} - y_{i-true}| + 1}}{n} \tag{1}$$

where n, y_{i-pred}, y_{i-true}, indicates the total number of samples, the i-th patient's severity level that the model predicts, and the actual severity level of the i-th patient, respectively.

3.3.2. Spatial analysis method

To explore the spatial agglomeration effect of the MI severity of in Wuhan, this paper uses the global Getis-Ord General G and local Getis-Ord Gi* methods (Getis and Ord, 2010). The local Getis-Ord Gi* is calculated as follows.

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \overline{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \tag{2}$$

where G_i^* denotes the z-score. Given that it is significant, the positive z-score means that the region is a hotspot, i.e., a high severity of MI exists in this region, and vice versa. x_j denotes the attribute value of j-th element; $w_{i,j}$ denotes the spatial weight between i-th element and j-th element; n denotes the total number of elements; $\overline{X}$ denotes the mean value of the attributes of all elements in a region; S denotes the standard deviation of the attributes of all elements in a region.

4. Results

4.1. General statistical analysis

A general statistical analysis is performed for the MI patient attendance data. Gender, age and temperature are found as the key factors influencing the MI severity. Older people and men are more likely to suffer a severe MI. And the mortality rate for the disease is higher in winter than in summer. The detailed results are presented in the

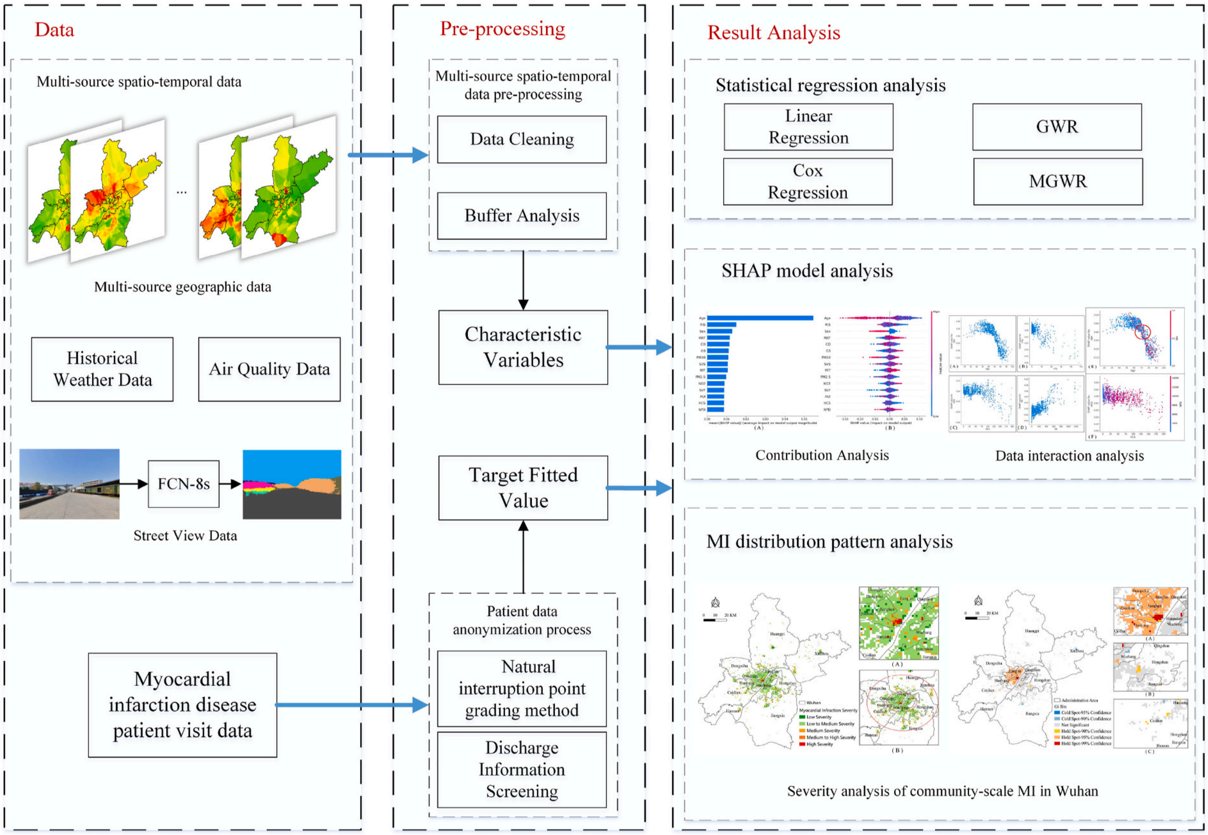

Fig. 2. Workflow of assessing MI severity from the urban environment perspective.

supplementary material Fig. S1 and Fig. S2.

4.2. Analysis of multi-source spatial factors for MI severity via RFA-based SHAP model

4.2.1. Performance of RFA-based SHAP model

This section demonstrates the effectiveness of the RFA for high-dimensional nonlinear problems. The confusion matrix of prediction of MI severity using the RFA-based model is shown in Table 2. The result shows a Kappa coefficient of 0.889, a Precision of 0.933, a Recall of 0.874 and an F1 score of 0.900. It can be seen that the RFA has a good performance in this issue.

We also set up a baseline experiment to demonstrate the good performance of the RFA. Table S2 shows that in the MI severity analysis, traditional regression models are not very accurate (R^2 is below 0.2) when faced with large amounts of non-collinear spatio-temporal data. In addition, the Cox frailty model (AIC = 1200.0, AUC = 0.956) fitted significantly better than MGWR (AIC = 3055.1) in the MI mortality analysis. While the RFA not only has the advantage of explaining the contribution of the multi-source variables, but also can fits the high-dimensional non-collinear data well (R^2 = 0.774 in the MI severity analysis, and AUC = 0.960 in the MI mortality analysis). Table S3 and Fig. S3 demonstrate the advantages of RFA over other machine learning algorithms. Moreover, the training set accuracy as well as the test set accuracy of the RFA are higher than other algorithms, further proving it is the optimal algorithm for this study.

4.2.2. The effects of multi-source spatial factors on MI severity and mortality

The effects of multi-source spatial factors on MI severity and mortality are analysed separately.

(1) In terms of MI severity, the risk factor that causes the risk of highest MI severity is age, followed by road intersections, patient gender, and meteorological condition, as shown in Fig. 3 (A). Fig. 3 (B) illustrates that (a) the MI severity is relatively higher for younger patients. (b) For residents living in areas with good access to transportation (RIS) and high quality of health care service

Table 2
Confusion matrix of classification. Columns represent predicted values, while rows represent ground-truth values.

Disease Severity	Low	Low-Medium	Medium	Medium-High	High
Low	264	35	4	1	1
Low-Medium	16	410	11	0	1
Medium	9	26	179	0	0
Medium-High	1	5	0	29	0
High	3	9	0	0	94

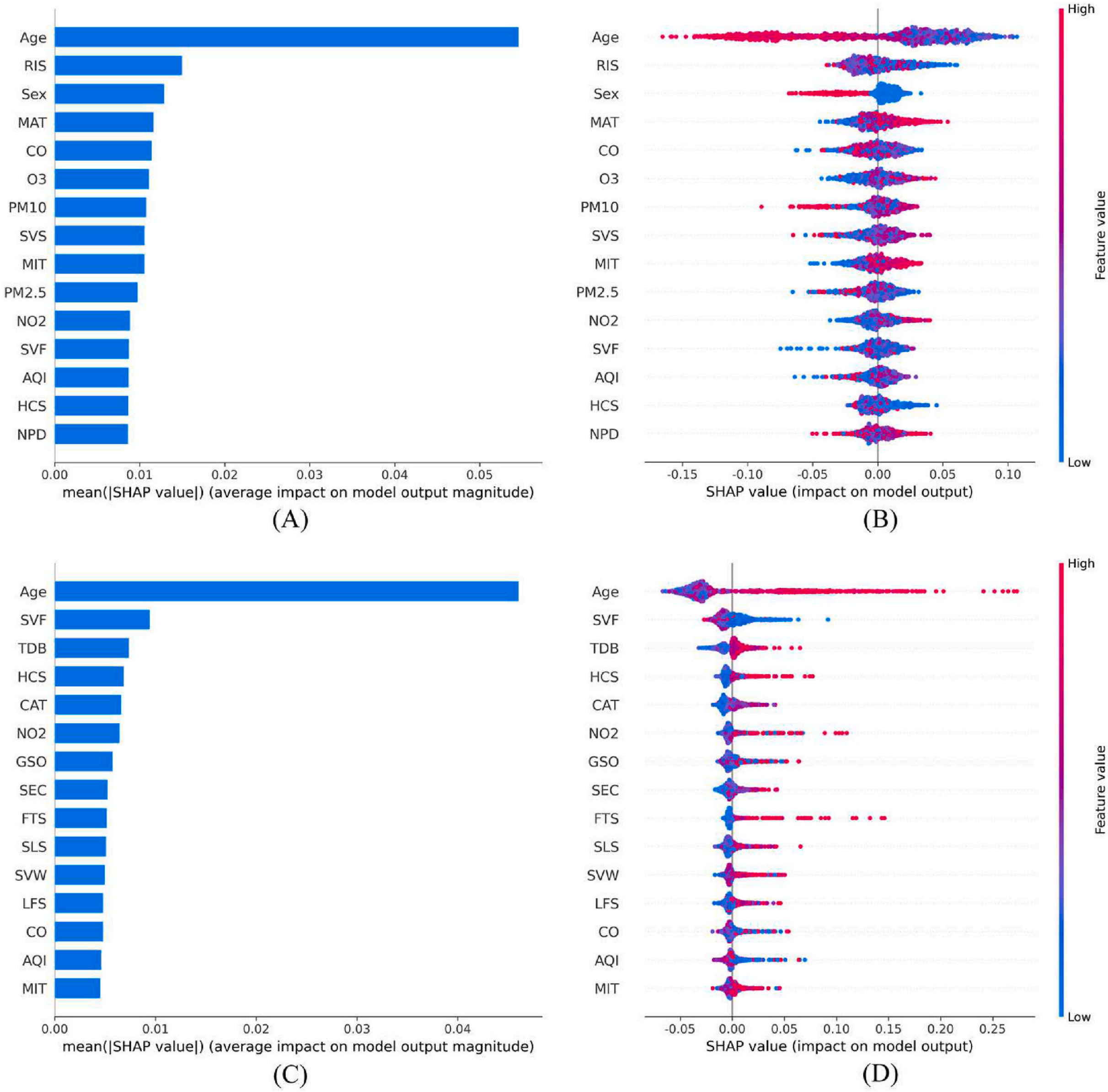

Fig. 3. Analysis of the risk factors for MI severity. (A) MI severity-variable contribution. (B) MI severity-variable-driven benefit. (C) MI mortality-variable contribution. (D) MI mortality-variable-driven benefit.

(HCS), they have the conditions for early detection and treatment of the disease. (c) Male patients are usually at a risk of significantly higher MI severity than female patients in terms of gender. And (d) residents are at a risk of higher MI severity of acute MI in both hot weather (MAT) and cold weather (MIT).

(2) In terms of MI mortality, age is the most important factor (Fig. 3 (C)). Fig. 3 (D) further found that (a) the older the infarction patient, the higher the mortality rate. (b) The mortality rate of MI patients is higher in colder temperatures (MIT), or in weather with a high temperature difference between days (TDB) (e.g., after a severe cooling). (c) The more health care services (HCS) are available, the more people in those areas choose to seek medical attention nearby. However, most general community hospitals are not equipped to treat MI, which delays the prime time for treatment and leads to a higher mortality rate.

The severity of MI varies by region and by population. As shown in Fig. 4 (A), residents living in scenic areas, such as scenic spots, are less likely to have an acute MI, suggesting that urban livability is important for residents' physical and mental health. Notably, Seo et al. (2019) found that increasing urban green space coverage would help reduce the risk of high CVD severity, based on the Cox proportional hazards regression model and the CVD records from over 300,000 Asian populations, confirming the importance of improving urban livability for population health (Seo et al. 2019). From Fig. 4 (D), it can be seen that residents living in suburban areas with low population density at night have a higher MI mortality rate as they take longer to reach a secondary or tertiary care hospital, delaying the prime time for treatment. Fig. 4 (B) and (E) show that the MI severity is lower in well-developed care services areas, but the MI mortality rate is higher in these areas. Fig. 4 (C) and (F) indicate that the MI severity is significantly higher for residents under 75 years of age than for those over 75 years of age. But the mortality rate is higher for older MI patients (over 75 years of age). The MI severity and mortality rates are significantly higher for women over 75 years of age. A previous study has also shown that older women are more likely to suffer a severe MI in bad weather (Li et al. 2019).

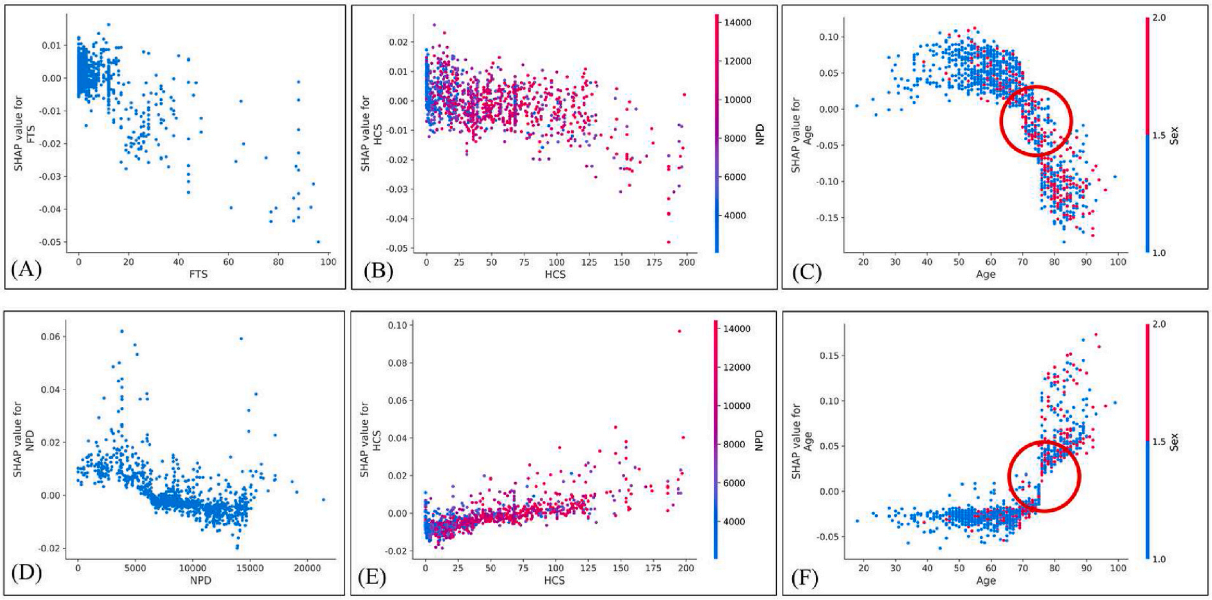

Fig. 4. MI severity driver interaction analysis, including (A) FTS-SHAP values, (B) HCS-NPD interaction analysis, (C) Age-Sex interaction analysis. Infarction mortality driver interaction analysis, i.e., (D) FTS-SHAP values, (E) HCS-NPD interaction analysis, (C) Age-Sex interaction analysis.

4.3. Community-scale spatial pattern of MI disease

We also analysed the spatial pattern of MI in Wuhan at a community scale. Fig. 5 (A) shows that the risk area for highest MI severity is located near the Jianghan Road pedestrian street in the central area of Hankou (Jiang'an District, Jianghan District and Qiaokou District). It has higher pedestrian and vehicular traffic flow, higher exposure to crowd activity and traffic noise, and poorer sleep quality, thereby raising the risk of higher acute MI severity. The MI severity is also higher in areas such as Wuchang District and Hongshan District, where amounts of universities are located. The MI severity is also higher in the suburb areas, showing a donut-shape pattern of medium-high severity, as shown in Fig. 5 (B).

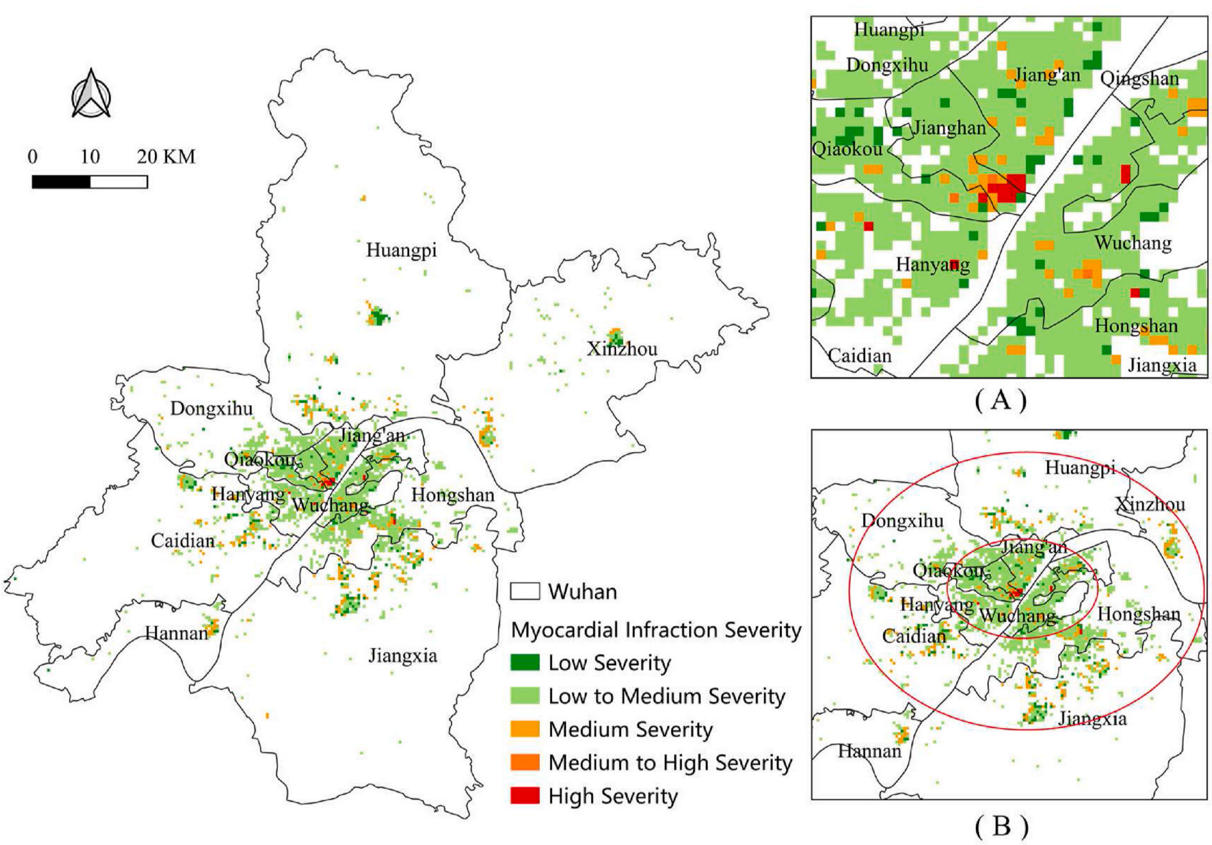

Fig. 5. Spatial distribution of MI severity at the community scale in Wuhan, (A) in the central city of Wuhan and (B) in satellite cities around the central city of Wuhan.

Among the medium-severity communities, 70.8% are in the suburb areas of Wuhan. This further confirms the finding in Fig. 4 (D) that in areas with poor accessibility, failure to seek medical attention in a timely manner will lead to increased severity of the disease.

The global Getis-Ord General G spatial analysis was conducted on the MI severity in Wuhan communities. A global G index of 0.002 is obtained, with a p of 0.001 (significant) and a z score of 3.265 (greater than 0). This indicates that overall in Wuhan, the MI severity is clustered in communities with high values, i.e., the distribution of higher severity communities is clustered.

In addition, the local Getis-Ord Gi* hotspot analysis of the MI severity in each community in Wuhan was implemented. As seen in Fig. 6 (A), the hotspot communities are mainly located in Jianghan District, Jiang'an District and Hanyang District with 99% confidence. It is related to the fact that this area is an economically developed urban old city in Wuhan with a large elderly population living in the city. Fig. 6 (B) reveals that hotspots with a confidence of 95% also exist in Hongshan District and Wuchang District, which are areas with a high concentration of universities in Wuhan. Fig. 6 (C) shows that a few hotspots with a confidence greater than or equal to 90% for the MI severity also exist in the suburbs of Wuhan (Caidian District), which is related to the poor configuration of medical facilities in the suburbs of the city and the lack of access to high-quality medical care for residents.

5. Discussion

5.1. Scientific contributions

This paper makes the following 3 scientific contributions.

First, this paper is the first to apply multi-source spatio-temporal data and RFA-SHAP model to analyse the risk factors for MI severity. We revealed that the RFA-based SHAP interpretable model is the best performing machine-learning model to analyse the quantitative impact of spatial risk factors on MI disease by comparing different baseline methods. And this paper innovatively predicts the potential spatial MI severity at the community scale, and analyses its abnormal spatial distribution. The results can be used for in-depth research in the fields of public health, big data and urban planning, and provide reasonable support for the targeted distribution of medical facilities at the community scale.

Second, this paper calls for attention to be paid to the impact of spatial environmental factors on MI severity. We discovered some interesting spatial risk factors and spatial patterns of MI severity and mortality, which can often be overlooked but are helpful for severe MI prevention and health equality in urban planning. Some spatial risk factors (e.g., areas with good access to medical and traffic services, pleasant scenery and high air quality) have a lower severity of residents suffering from MI, which is consistent with previous studies (Chen et al. 2020; Widimský et al. 2003). However, we found that the mortality is higher in areas with too many medical services (50+ homes/600 m buffer zone). Also, a higher mortality of acute MI is found in extreme weather conditions (e.g., low temperatures and sudden temperature changes). Spatial patterns of MI at the community scale were discovered for urban planning in Wuhan, China. We found that the MI severity is high in the vicinity of Jianghan Road Pedestrian Street in Jianghan District, where are the traditional downtown area of Wuhan. While the MI severity is high in the suburb areas, with a strong spatial agglomeration effect.

Last, this paper summarised some rules of MI severity for urban residents in terms of age and gender. Residents who meet the following conditions should pay more attention to the prevention of severe MI. (a) In terms of patient age, younger patients tend to have a higher MI

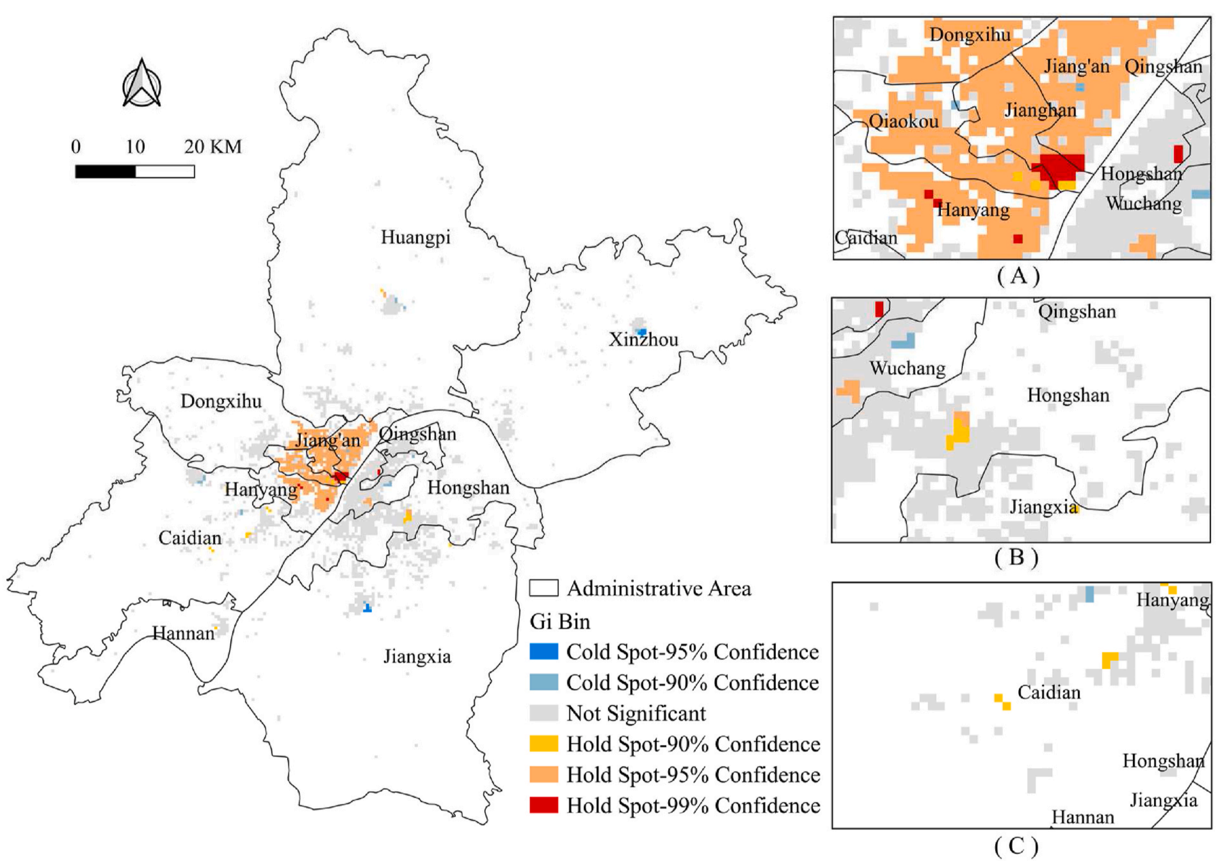

Fig. 6. Analysis of the distribution pattern of MI severity hotspots at the district scale in Wuhan, (A) in the central city of Wuhan, (B) in Wuchang and Hongshan districts of Wuhan, and (C) in Caidian district of Wuhan.

severity but older patients have a higher mortality rate. 75 years of age is the cut-off point. Approximately 5% of those over 75 years of age presents with symptoms of MI without a history (Valensi et al. 2011). (b) In terms of patient gender, women with MI have a higher average age and mortality rate than men, while men have a higher MI severity than women. The average age of female patients in hospital is 73.57 years and the mortality rate is 12.06%, while the average age of male patients is 64.13 years and the mortality rate is 8.46%.

5.2. Policy implication

Based on the above results, the following three advices are made for the prevention of severe MI in the aspect of policy implication.

(a) Strengthen residents' education on severe MI prevention and first aid science. Because these knowledge about MI facilitates the timely treatment of acute MI cases (Hertz et al. 2019). This paper recommends equipping residents with basic acute MI countermeasures and being able to arrange routes to medical care as soon as possible, while concentrating on severe MI prevention measures for the elderly (aged 75 and above) in extreme weather conditions.
(b) Increase the prevalence of medical check-ups among the residents. Early detection and treatment of MI are essential to reduce the mortality from MI (Laichuthai et al. 2020), as the acute MI is associated with underlying conditions such as the patient's own hypertension (Flint et al. 2019) and diabetes (Norhammar et al. 2002). Thus, we suggest strengthening regular CVDs screening for residents in some areas, e.g., densely populated communities, colleges and universities, and areas with a large elderly population.
(c) Adhere to the principle of equality in urban planning. The increase in visible green space (Seo et al. 2019), the reduction of urban noise (Nieuwenhuijsen, 2018) and the equitable distribution of medical facilities (Miller et al. 2020) can help to reduce the incidence of MI among residents. Thus, this paper proposes to upgrade the greening of the central city, build road sound barriers in the downtown area to reduce urban traffic noise, and rationally distribute good-quality medical facilities in both urban and suburb areas.

5.3. Shortcomings and future work

Inevitably, some shortcomings exist in this study.

(1) This paper finds that the SHAP model can perfectly integrate advanced machine learning algorithms and apply its excellent interpretability to the field of public health as a bold attempt. However, machine learning models require a large data size for training, otherwise the models are prone to overfitting (Breiman, 2001). It is well known that data are challenging to obtain in the medical field. Thus, the next step in this research is to get better results with small data.
(2) The RFA-SHAP has a longer computation time, compared to the conventional machine learning models (e.g., multiple linear regression, Cox frailty model and GWR). It is also worth considering how to improve the computational speed of the model based on high performance computing technology.
(3) Only information of age and gender were collected. Other body factors and health issues (e.g., BMI, blood pressure, drinking/smoking habit, obesity, diabetes, exercise habit, family history, etc.) are also important to be considered. In future studies, we will further analyse patient attendance data and field survey data at the community scale rather than the individual patient scale. Furthermore, this study will carry out a more comprehensive multi-dimensional analysis of the risk factors for comparison.

6. Conclusion

This paper proposes the RFA-based SHAP interpretable model to explore the impact of spatial risk factors on MI severity. The model shows a good performance with the test accuracy of 0.720 and the Kappa coefficient of 0.889. In addition, we found that.

(1) strong correlations exist between MI severity and age, gender of patients. First, younger patients tend to get a higher severity of MI. Second older patients show a higher MI mortality rate, especially those over 75 years of age. Next, males are at higher severity of getting MI than females. Last, female patients have a higher mortality rate than males.
(2) MI severity is related to living environment and meteorology factors, which can often be ignored. First, residents in areas with good access to medical facilities and high urban livability are at risk of lower severity of MI. Second, acute MI mortality is higher in areas with too many medical facilities. Next, extreme weather conditions such as low temperatures and sudden temperature changes will also increase the mortality rate of MI.
(3) Some spatial patterns of MI severity are revealed in Wuhan. First, residents in downtown areas of urban centres are more likely to get high-severity MI. Second, in suburb areas where medical and traffic facilities are not well-equipped, a donut-shape pattern of risk of medium-to-high-severity of MI exists. This paper aims to appeal for attention to the impact of environmental risk factors on MI severity. Thus, three recommendations are made to reduce the spatial impact on MI severity in the aspect of policy implication. We propose (a) popular education of residents on MI prevention, (b) regular medical check-ups for community residents, and (c) rational allocation of city resources to ensure health equality within the city.

Author contribution statement

Prof. Yao Yao: Conceptualization Ideas, Methodology, Formal analysis, Writing – original draft, Funding acquisition. Mr. Hanyu Yin: Software Programming, Investigation. Dr. Changwu Xu: Data curation, Validation, Resources. Mr. Dongsheng Chen: Writing – review & editing, Formal analysis. Mr. Ledi Shao: Visualization, Formal analysis. Prof. Qingfeng Guan: Project administration, Resources, Editing. Dr. Ruoyu Wang: Resources, Editing.

Fundings

This work was supported by the National Key Research and Development Program of China (Grant No. 2019YFB2102903), the National Natural Science Foundation of China (Grant No. 41801306), the "CUG Scholar" Scientific Research Funds at China University of Geosciences (Wuhan) (Project No. 2022034).

Declaration of competing interest

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

Acknowledgements

This work was supported by the National Key Research and Development Program of China (Grant No. 2019YFB2102903), the National Natural Science Foundation of China (Grant No. 41801306), the "CUG Scholar" Scientific Research Funds at China University of Geosciences (Wuhan) (Project No. 2022034).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2022.115438.

References

Akbarzadeh, M.A., et al., 2018. The association between exposure to air pollutants including PM10, PM2. 5, ozone, carbon monoxide, sulfur dioxide, and nitrogen dioxide concentration and the relative risk of developing STEMI: a case-crossover design. Environ. Res. 161, 299–303.

Amsalu, E., et al., 2019. Spatial-temporal analysis of cause-specific cardiovascular hospital admission in Beijing, China. Int. J. Environ. Health Res. 1–12.

Asri, A.K., et al., 2020. Global greenness in relation to reducing the burden of cardiovascular diseases: ischemic heart disease and stroke. Environ. Res. Lett. 15 (12), 124003.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Chen, H., et al., 2020. Residential greenness and cardiovascular disease incidence, readmission, and mortality. Environ. Health Perspect. 128 (8), 87005.

Chen, J., et al., 2013. Research on geographical environment unit division based on the method of natural breaks (Jenks). Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. 3, 47–50.

Chen, K., et al., 2019. Temporal variations in the triggering of myocardial infarction by air temperature in Augsburg, Germany, 1987–2014. Eur. Heart J. 40 (20), 1600–1608.

Cox, D.R., Wermuth, N., 1992. A comment on the coefficient of determination for binary responses. Am. Statistician 46 (1), 1–4.

Flach, P.A., Kull, M., 2015. Precision-Recall-Gain Curves: PR Analysis Done Right. NIPS.

Flint, A.C., et al., 2019. Effect of systolic and diastolic blood pressure on cardiovascular outcomes. N. Engl. J. Med. 381 (3), 243–251.

Fotheringham, A.S., Yang, W., Kang, W., 2017. Multiscale geographically weighted regression (MGWR). Ann. Assoc. Am. Geogr. 107 (6), 1247–1265.

Franch-Pardo, I., et al., 2020. Spatial analysis and GIS in the study of COVID-19. A review. Sci. Total Environ. 739, 140033.

Gerber, Y., et al., 2010. Neighborhood socioeconomic context and long-term survival after myocardial infarction. Circulation 121 (3), 375–383.

Getis, A., Ord, J.K., 2010. The analysis of spatial association by use of distance statistics. In: *Perspectives on Spatial Data analysis*Springer), pp. 127–145.

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. Geojournal 69 (4), 211–221.

Helbich, M., et al., 2019. Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. Environ. Int. 126, 107–117.

Hertz, J.T., et al., 2019. Knowledge of myocardial infarction symptoms and perceptions of self-risk in Tanzania. Am. Heart J. 210, 69–74.

Hu, Z., Liebens, J., Rao, K.R., 2008. Linking stroke mortality with air pollution, income, and greenness in northwest Florida: an ecological geographical study. Int. J. Health Geogr. 7 (1), 1–22.

Kang, K., Wang, X., 2014. Fully Convolutional Neural Networks for Crowd Segmentation, 1411.4464.

Kihal-Talantikite, W., et al., 2017. Developing a data-driven spatial approach to assessment of neighbourhood influences on the spatial distribution of myocardial infarction. Int. J. Health Geogr. 16 (1), 1–18.

Lachowycz, K., Jones, A.P., 2014. Does walking explain associations between access to greenspace and lower mortality? Soc. Sci. Med. 107, 9–17.

Laichuthai, N., et al., 2020. Newly discovered abnormal glucose tolerance in patients with acute myocardial infarction and cardiovascular outcomes: a meta-analysis. Diabetes Care 43 (8), 1958–1966.

Li, J., et al., 2019. Association between ambient particulate matter air pollution and ST-elevation myocardial infarction: a case-crossover study in a Chinese city. Chemosphere 219, 724–729.

Lin, Y., Jeon, Y., 2006. Random forests and adaptive nearest neighbors. J. Am. Stat. Assoc. 101 (474), 578–590.

Liu, H., et al., 2018. Fine particulate air pollution and hospital admissions and readmissions for acute myocardial infarction in 26 Chinese cities. Chemosphere 192, 282–288.

Lundberg, S.M., Lee, S., 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 4768–4777.

Maclure, M., Mittleman, A.M., 2000. Should we use a case-crossover design? Annu. Rev. Publ. Health 21 (1), 193–221.

Members, A.T.F., et al., 2012. ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: the Task Force on the management of ST-segment elevation acute myocardial infarction of the European Society of Cardiology (ESC). Eur. Heart J. 33 (20), 2569–2619.

Miller, I.F., et al., 2020. Disease and healthcare burden of COVID-19 in the United States. Nat. Med. 26 (8), 1212–1217.

Miller, K.A., et al., 2007. Long-term exposure to air pollution and incidence of cardiovascular events in women. N. Engl. J. Med. 356 (5), 447–458.

Mitchell, R., Popham, F., 2008. Effect of exposure to natural environment on health inequalities: an observational population study. The lancet 372 (9650), 1655–1660.

Mostofsky, E., et al., 2015. Risk of myocardial infarction immediately after alcohol consumption. Epidemiology 26 (2), 143.

Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4 (2), 133–142.

Nawrot, T.S., et al., 2011. Public health importance of triggers of myocardial infarction: a comparative risk assessment. Lancet 377 (9767), 732–740.

Nieuwenhuijsen, M.J., 2018. Influence of urban and transport planning and the city environment on cardiovascular disease. Nat. Rev. Cardiol. 15 (7), 432–438.

Norhammar, A., et al., 2002. Glucose metabolism in patients with acute myocardial infarction and no previous diagnosis of diabetes mellitus: a prospective study. Lancet 359 (9324), 2140–2144.

Raziani, Y., Raziani, S., 2021. The effect of air pollution on myocardial infarction. J. Chem. Rev. 3 (1), 83–96.

Richardson, E., et al., 2010. The association between green space and cause-specific mortality in urban New Zealand: an ecological analysis of green space utility. BMC Publ. Health 10 (1), 1–14.

Richardson, E.A., et al., 2012. Green cities and health: a question of scale? J. Epidemiol. Community Health 66 (2), 160–165.

Richardson, E.A., Mitchell, R., 2010. Gender differences in relationships between urban green space and health in the United Kingdom. Soc. Sci. Med. 71 (3), 568–575.

Rowland, S.T., et al., 2020. Can ultra short-term changes in ambient temperature trigger myocardial infarction? Environ. Int. 143, 105910.

Rubinstein, R.Y., Kroese, D.P., 2016. Simulation and the Monte Carlo Method. John Wiley & Sons.

Rushton, G., 2003. Public health, GIS, and spatial analytic tools. Annu. Rev. Publ. Health 24 (1), 43–56.

Seo, S., et al., 2019. Association between urban green space and the risk of cardiovascular disease: a longitudinal study in seven Korean metropolitan areas. Environ. Int. 125, 51–57.

Smyth, A., et al., 2016. Physical activity and anger or emotional upset as triggers of acute myocardial infarction: the INTERHEART study. Circulation 134 (15), 1059–1067.

Statistics, N.B.O., 2020. Wuhan Statistical Yearbook. China Statistical Publishing House (4), pp. 88–91.

Tamosiunas, A., et al., 2014. Accessibility and use of urban green spaces, and cardiovascular health: findings from a Kaunas cohort study. Environ. Health 13 (1), 1–11.

Timonin, S., et al., 2018. Reducing geographic inequalities in access times for acute treatment of myocardial infarction in a large country: the example of Russia. Int. J. Epidemiol. 47 (5), 1594–1602.

Tonne, C., et al., 2007. A case–control analysis of exposure to traffic and acute myocardial infarction. Environ. Health Perspect. 115 (1), 53–57.

Valensi, P., Lorgis, L., Cottin, Y., 2011. Prevalence, incidence, predictive factors and prognosis of silent myocardial infarction: a review of the literature. Arch. Cardiovascul. Dis. 104 (3), 178–188.

Vienneau, D., et al., 2019. Façades, floors and maps–Influence of exposure measurement error on the association between transportation noise and myocardial infarction. Environ. Int. 123, 399–406.

Villeneuve, P.J., et al., 2012. A cohort study relating urban green space with mortality in Ontario, Canada. Environ. Res. 115, 51–58.

Vos, T., et al., 2016. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The lancet 388 (10053), 1545–1602.

Wang, R., et al., 2019. Cross-sectional associations between long-term exposure to particulate matter and depression in China: the mediating effects of sunlight, physical activity, and neighborly reciprocity. J. Affect. Disord. 249, 8–14.

Wang, R., et al., 2020. Relationship between eye-level greenness and cycling frequency around metro stations in Shenzhen, China: a big data approach. Sustain. Cities Soc. 59, 102201.

White, H.D., Chew, D.P., 2008. Acute myocardial infarction. Lancet 372 (9638), 570–584.

Widimský, P., et al., 2003. Long distance transport for primary angioplasty vs immediate thrombolysis in acute myocardial infarction: final results of the randomized national multicentre trial—PRAGUE-2. Eur. Heart J. 24 (1), 94–104.

Xu, Y., et al., 2018. Association between the metabolically healthy obese phenotype and the risk of myocardial infarction: results from the Kailuan study. Eur. J. Endocrinol. 179 (6), 343–352.

Yang, B., et al., 2020. Association between residential greenness, cardiometabolic disorders, and cardiovascular disease among adults in China. JAMA Netw. Open 3 (9), e2017507.

Yang, J., et al., 2017. Seasonal variations of temperature-related mortality burden from cardiovascular disease and myocardial infarction in China. Environ. Pollut. 224, 400–406.

Yao, Y., et al., 2017a. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. Int. J. Geogr. Inf. Sci. 31 (4), 825–848.

Yao, Y., et al., 2017b. Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. Int. J. Geogr. Inf. Sci. 31 (6), 1220–1244.

Yao, Y., et al., 2019. A human-machine adversarial scoring framework for urban perception assessment using street-view images. Int. J. Geogr. Inf. Sci. 33 (12), 2363–2384.

Yao, Y., et al., 2021. Transmission risk assessment of novel coronavirus pneumonia epidemic in Wuhan city habitat based on fine scale of multi-source geographical data. Acta Ecol. Sin. 41 (19), 7493–7508.

Yao, Y., et al., 2022. More visible greenspace, stronger heart? Evidence from ischaemic heart disease emergency department visits by middle-aged and older adults in Hubei, China. Landsc. Urban Plann. 224, 104444.

Yuan, N.J., et al., 2014. Discovering urban functional zones using latent activity trajectories. IEEE Trans. Knowl. Data Eng. 27 (3), 712–725.