



# Evaluating Spatial Accessibility to COVID-19 Vaccination Sites based on Fine-Scale Population Distributions and Heterogeneous Travel Modes: A Case Study in Xiangtan, China

Wentao Yang<sup>1,2</sup> · Fengjie Wang<sup>1,2</sup> · Yihan You<sup>2</sup> · Xiafan Wan<sup>1,2</sup> · Sijie Cheng<sup>3</sup> · Zhixiong Fang<sup>4,5</sup>

Received: 1 August 2023 / Accepted: 18 December 2023  
© The Author(s) 2024

## Abstract

Understanding the inequity in spatial accessibility to COVID-19 vaccination sites is beneficial for the government to optimize vaccine resources. Nevertheless, the result of evaluating spatial accessibility to COVID-19 vaccination sites may be directly affected by the lack of fine-scale and accurate population distribution information in both urban and rural areas. In addition, the choice of travel modes related to different age groups has not been incorporated into existing methods well. Consequently, this study proposes a framework to evaluate spatial accessibility to COVID-19 vaccination sites based on fine-scale population distributions and heterogeneous travel modes. A divide-and-conquer strategy is first presented to estimate the fine-scale populations in both urban and rural areas. A hybrid time distance considering heterogeneous travel modes and a population-weighted proximity indicator are then defined to evaluate spatial accessibility for each age group. The experimental results from Xiangtan, China, show that (1) the divide-and-conquer strategy can effectively predict the urban and rural populations. (2) Similar clustering patterns of spatial accessibility are observed for each age group. (3) Inequities have been confirmed at the county level. The main findings in this study can provide valuable information supporting the spatial optimization of existing COVID-19 vaccination sites.

**Keywords** COVID-19 · Vaccination · Spatial analysis · Spatial accessibility · Inequity

## Introduction

The COVID-19 pandemic has posed an extremely serious and unprecedented challenge for governments worldwide (Anderson et al., 2020). During the early stage of the pandemic, nonpharmaceutical interventions such as school closures, travel

---

Extended author information available on the last page of the article

restrictions, and public lockdowns were commonly used to slow or pause the spread of the disease (Qi et al., 2022). Due to various socioeconomic factors and policies across different countries or regions, it is difficult to strictly and continuously implement such interventions, which greatly affects the efficiency of global epidemic prevention and control (Duhon et al., 2021; Kaufman et al., 2021). Fortunately, with international vaccine development efforts, COVID-19 vaccines have been developed, and vaccination has been proven to be a highly effective tool to fight the global COVID-19 pandemic (Bucyibaruta et al., 2022). Nevertheless, because of the uneven distribution of COVID-19 vaccination sites, variation in spatial accessibility to vaccination sites always exists for different groups (Liu et al., 2022). Understanding the inequity in spatial accessibility to vaccination sites is beneficial for governments to optimize vaccine resources.

Spatial accessibility refers to the ability of residents to reach their destination or infrastructure and indicates the convenience of socioeconomic activities (Kwan et al., 2003; Tao et al., 2020b). Here, the destination or infrastructure is COVID-19 vaccination sites; that is, spatial accessibility to COVID-19 vaccination sites describes residents' ability to reach COVID-19 vaccination sites. A series of studies have been carried out to analyse spatial accessibility to vaccination sites, which can be distinguished from two aspects: one is the spatial scale, and the other is the measurement method of spatial accessibility. Spatial scales, the sizes or resolutions of spatial analysis units, heavily depend on the granularity of the initial observations or statistical values (Deng et al., 2017). Official census data are commonly employed to analyse spatial accessibility to COVID-19 vaccination sites (Qi et al., 2022; Tao et al., 2020a, b). Due to public privacy protection, census data are aggregated and released at the subdistrict or street scale (Bakillah et al., 2014), while the spatial analysis unit generally corresponds to the census tract level. However, it is difficult to accurately obtain the geographical location of residents according to census data. Global population products, such as the Gridded Population of the World (GPW), provide detailed population distribution information at a fine scale (Doxsey-Whitfield et al., 2015). These products based on the global scale are suitable for global or regional analysis, but due to local accuracy deviation, they are not suitable for local or fine-scale analysis (Yang et al., 2023). For this purpose, accurately obtaining fine-scale population distributions from census data is still challenging.

Recently, multiple sources of geographical data, such as mobile phone data, social media data, satellite-derived nighttime image data, and points of interest (POIs), have been proven to provide effective support for accurately predicting population distributions (Yao et al., 2017; Zeng et al., 2011). The theoretical assumption is that population density can be indirectly described by human activity intensity, and places with high activity intensity usually correspond to those with high population density. Based on these different types of geographical data, statistical regression or machine learning methods have been widely used to realize census data spatialization (Liu et al., 2018; Zhou et al., 2020). The essence of both statistical regression and machine learning methods tends to approximate the relationships between the population density and its associated variables of human activity. However, due to the differences in lifestyle and production modes, the spatial distribution of urban and rural populations presents different characteristics. The urban

population is mainly concentrated in urban built-up areas, while the rural population is relatively scattered. The geographical environment and facilities in urban and rural areas are also different. Geographical data for recording human activity, such as POIs and social media data, may be seriously lacking in rural areas (Liu et al., 2020). Obviously, because of the different characteristics representing human activity, existing methods for predicting the fine-scale population in urban areas cannot be used to analyse the population in rural areas. It is difficult to comprehensively evaluate the inequity in spatial accessibility to vaccination sites based only on the urban population and without considering the rural population. Therefore, how to generate the fine-scale population distribution of urban and rural areas is an urgent problem that needs to be solved.

Like other infrastructures, the measurement method of spatial accessibility to vaccination sites can be roughly divided into two categories, namely, distance-based indicators and spatial interaction methods (Kwan et al., 2003; Xiao et al., 2022). Distance-based indicators aim to measure spatial accessibility by the distance (such as space, time, and cost) between the residents' location and the vaccination sites. For instance, Guhlincozzi and Lotfata (2021) used the street-network distance to evaluate spatial accessibility to vaccination sites in the metropolitan area of Chicago, Illinois, United States. The street-network distance can measure the actual moving distance of a resident and is more reasonable than the Euclidean distance. Overall, it is easy to implement spatial accessibility by distance-based indicators, but they cannot effectively describe the relationships between spatial supply and demand. Spatial interaction methods tend to consider the spatial variation in vaccine supply and demand based on travel distance. These methods mainly include the gravity-based method and the two-step floating catchment area (2SFCA) method. For instance, Liu et al. (2022) explored racial/ethnic inequity in spatial accessibility to COVID-19 vaccination sites using a gravity-based model. Qi et al. (2022) used the 2SFCA method to investigate spatial accessibility to COVID-19 vaccination sites in two counties in the United States. Although spatial supply and demand can be addressed by these spatial interaction methods, how to measure and match the supply and demand needs to be further discussed.

Actually, both distance-based indicators and spatial interaction methods have their own advantages and limitations, and neither can completely replace the other. The kind of method that should be chosen depends on the application scenario. Specifically, during the initial vaccination stage, due to the limitation of vaccine production capacity, there is an imbalance between the supply and demand of vaccines. In this scenario, where the spatially varying supply and demand of vaccines should be considered, spatial interaction models are more suitable than distance-based indicators. However, with the rapid improvement of vaccine production capacity, the vaccine supply is sufficient in many countries or regions. In this scenario, distance-based indicators are more effective than spatial interaction methods.

Whether the supply and demand of vaccines should be considered in the method for measuring spatial accessibility to COVID-19 vaccination sites or not, the distances between the locations of residents and vaccination sites should be determined by both methods. A variety of distance measurement strategies have been proposed. Street-network distance, a commonly used strategy, is easy to implement

(Qi et al., 2022), and it may be applicable for people with disabilities and older people who general use a single travel mode (Guhlincozzi & Lofata, 2021). However, when different travel modes are involved, the single street-network distance cannot reveal the impact of travel modes on spatial accessibility. To overcome this limitation, Tao et al. (2020a, b) proposed that two primary travel modes, i.e., walking and driving, should be incorporated into the 2SFCA method, and car-accessible and car-less populations were distinguished to determine different distance thresholds and demand capacities for different travel modes. In fact, the choice of travel mode mainly depends on the travel distance or time between the origin and destination. The closer the origin and destination are, the more likely the walking mode will be chosen, and conversely, the more likely the driving mode will be selected. The choice of travel mode is also related to different age groups. How to incorporate heterogeneous travel modes into the spatial accessibility to COVID-19 vaccination sites is another urgent issue to be explored.

Consequently, the purpose of this study is to evaluate spatial accessibility to COVID-19 vaccination sites based on fine-scale population distributions and heterogeneous travel modes. The main objectives are (1) to propose a divide-and-conquer strategy to predict the fine-scale population distributions in urban and rural areas to determine the fine-scale demand capacity of COVID-19 vaccination sites; (2) to construct a distance-based indicator with the integration of heterogeneous travel modes to evaluate spatial accessibility to COVID-19 vaccination sites; and (3) to explore the inequity in spatial accessibility to COVID-19 vaccination sites in the study area.

The structure of this paper is organized as follows. In Section "Study area and Datasets", the study area and datasets are first introduced. The methods for predicting fine-scale population distributions and evaluating spatial accessibility are provided in Section "Methods". In Section "Results", the experimental results in the study area are described. The discussion and conclusion are provided in Sections "Discussion" and "Conclusions", respectively.

## Study area and Datasets

### Study Area

The study area, Xiangtan, is located in the east-central part of Hunan Province, China. The prefecture-level city consists of two districts (Yuhu and Yuetang) and three county-level cities (Xiangtan, Shaoshan, and Xiangxiang counties). The study area covers a total of 5,005.8 km<sup>2</sup>. The Seventh National Census of China reported that the permanent population in Xiangtan was 2,726,181, 50.74% male and 49.26% female, in 2020. The urban population was 1,754,969, while the rural population was 971,212. The spatial locations of the study area are shown in Fig. 1. Existing research has proven that the spatial distribution of medical services in the study area is unreasonable at the county level (Li et al., 2018); therefore, this research aims to evaluate spatial accessibility to COVID-19 vaccination sites at a fine spatial scale.

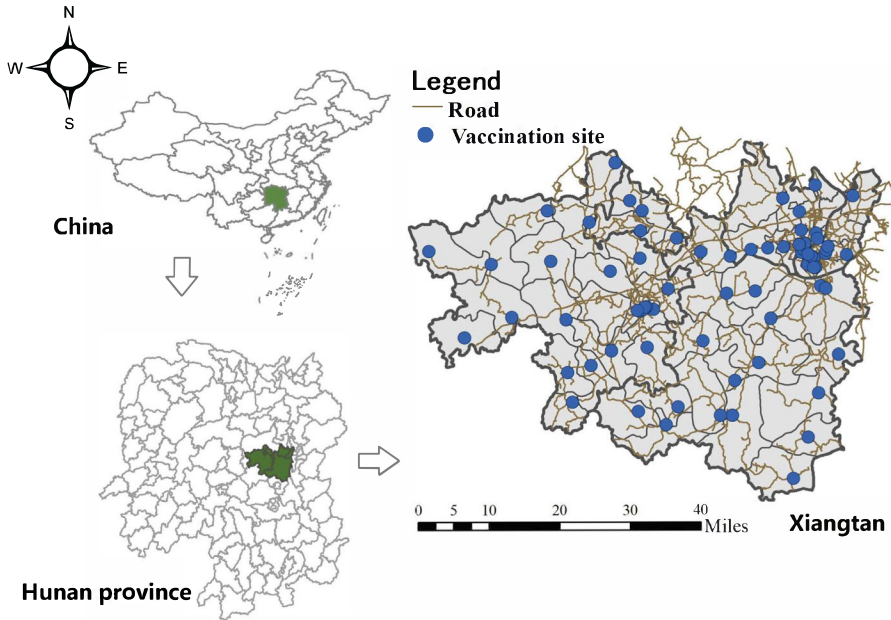


Fig. 1 Map of the study area

## Datasets and Preprocessing

The datasets used in this study mainly consist of nine types of data, which are listed as follows: (1) Census data of 70 units at the street scale in Xiangtan from the census statistical yearbook, including the population, the age structure information of each street, and the type of each street (urban or rural areas); (2) COVID-19 vaccination site data, including names and locations, which were derived from the Xiangtan Center for Disease Control and Prevention. The spatial distribution of COVID-19 vaccination sites is shown in Fig. 1; (3) Digital elevation model (DEM) data with a spatial resolution of 30 m from the Computer Network Information Center of the Chinese Academy of Sciences; (4) Fundamental geographical information, namely, the geometric information of the study area, which was obtained from National Geomatics Center of China; (5) High-resolution remote sensing images from the Gaofen-2 satellite with a spatial resolution of 5 m, provided from China Centre for Resources Satellite Data and Application; (6) Land use data (e.g., construction land, garden land, grassland, forestland, cultivated land, and other categories) with a spatial resolution of 30 m obtained from the Computer Network Information Center of the Chinese Academy of Sciences; (7) Nighttime light remote sensing images of Luojiang1-01 with a spatial resolution of 130 m; (8) Points of interest (POIs) data (public services points, business points, industrial points, residential points, and transportation points) with approximately 200,000 records that were obtained from Baidu maps; (9) Road network data obtained from OpenStreetMap; (10) Survey about travel

mode selection derived from a questionnaire (“age”, “how long are you willing to walk to get vaccinated?”), and approximately 100 samples were obtained for each age group (children, middle-aged people, and older people); (11) Urban statistical yearbook, including the total population and the numbers of the three age groups (children less than 16 years of age, middle-aged people from 16 to 60 years of age, and older people of more than 60 years of age) for each street, which was published by the Hunan Provincial Bureau of Statistics; and (12) Walking trajectories of different age groups, with a total sample size of 100. The details of these multisource datasets are listed in Table 1.

This study involves two types of basic analytical units: street units and grid units. Street units are known and obtained from fundamental geographical information data, but the resolution of grid units needs to be specified. The resolution of the grid units was set to 1 km, and the multiscale issue was not discussed in depth. In most existing studies about population spatialization, the street units were regarded as the initial units for building the relationships between population density and the covariables. Then, the built model was used to predict the population density of each grid unit. During data preprocessing, the covariable data were aggregated into street and grid units. The areas of these five types of land use, the numbers of the five POI categories, the average slope, the road network density, and the nighttime light intensity were calculated for each type of spatial unit based on the datasets. Hence, the covariables or input features are composed of the areas of the five types of land use, the average nighttime light intensity, the average slope, the numbers of the five POI categories, and the road network density.

**Table 1** Multisource datasets used in this study

Name	Year	Spatial scale (Data type)	Source
Census	2020	Street scale	<a href="http://xtj.xiangtan.gov.cn/">http://xtj.xiangtan.gov.cn/</a>
COVID-19 vaccination sites	2021	(points)	<a href="http://www.xtcdc.org.cn/">http://www.xtcdc.org.cn/</a>
DEM	2017	30 m (grid)	<a href="http://www.gscloud.cn">http://www.gscloud.cn</a>
Fundamental geographical information	2015	(Vector map)	<a href="http://www.ngcc.cn/ngcc/">http://www.ngcc.cn/ngcc/</a>
Land use	2018	30 m (grid)	<a href="https://www.resdc.cn">https://www.resdc.cn</a>
High-resolution remote sensing images	2018	5 m (grid)	<a href="https://www.cresda.com/zgzywxzyzx/index.html">https://www.cresda.com/zgzywxzyzx/index.html</a>
Nighttime light remote sensing images	2018	130 m (grid)	<a href="http://59.175.109.173:8888">http://59.175.109.173:8888</a>
POIs	2021	(Points)	<a href="http://map.baidu.com">http://map.baidu.com</a>
Road network	2021	(Networks)	<a href="https://www.openstreetmap.org">https://www.openstreetmap.org</a>
Survey about travel modes selection	2021	(Table)	Survey based on questionnaire
Urban statistical yearbook	2018	(Table)	<a href="http://tjj.hunan.gov.cn/tjsj/tjnj/">http://tjj.hunan.gov.cn/tjsj/tjnj/</a>
Walking trajectories of different age groups	2021	(Trajectory)	Manual observation

## Methods

The framework for evaluating spatial accessibility to COVID-19 vaccination sites based on fine-scale population distributions and heterogeneous travel modes is composed of three levels, namely, the task level, the method level, and the data level, as shown in Fig. 2. The task level mainly involves two aspects: one is to predict the fine-scale population, and the other is to measure spatial accessibility to COVID-19 vaccination sites. To accomplish these two tasks, a divide-and-conquer strategy and a distance-based indicator considering heterogeneous travel modes are developed based on the datasets described in Section "Study area and Datasets". The data needed for each level and method are shown in Fig. 2, and the principles of these two methods are introduced in detail in Sections "A Divide-and-Conquer Strategy for Predicting Fine-Scale Population Distributions" and "A Distance-based Indicator Considering Heterogeneous Travel Modes for Measuring Spatial Accessibility".

### A Divide-and-Conquer Strategy for Predicting Fine-Scale Population Distributions

As described in Section "Introduction", urban and rural populations always exhibit different distribution patterns, so different data and models are required to model the urban and rural population distributions at a fine scale. Consequently, a divide-and-conquer strategy is proposed to predict the fine-scale population distribution. The study area is first divided into two parts: urban and rural areas. For urban areas, a two-level random forest model is employed to predict the fine-scale population distribution (Yang et al., 2023). For rural areas, U-Net convolutional neural networks are first used to extract building image objects from high-resolution remote sensing images (Kim et al., 2019), and the areas of building image objects are then defined as an allocation index to allocate the population.

Recently, machine learning methods have been widely used to predict the population distribution based on new forms of ancillary data, such as POIs and nighttime light remote sensing images (Jia & Gaughan, 2016). Compared with spatial statistical methods, these methods can adaptively learn any nonlinear relationships, especially the random forest model, which has been proven to be outstanding in describing the associations between population density and covariables. In

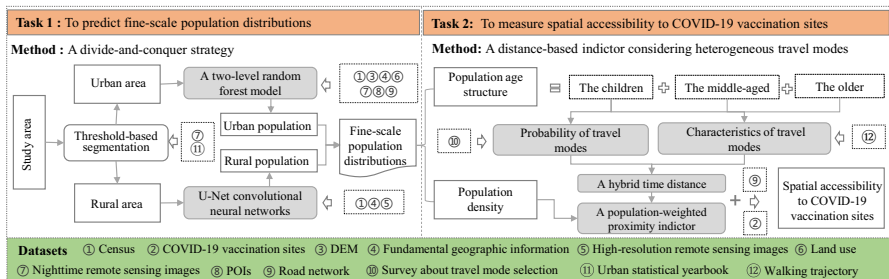


Fig. 2 Framework for evaluating spatial accessibility in this study

this research, a variant of the random forest model, namely, the two-level random forest model, was selected to predict the population distribution in the urban area (Yang et al., 2023).

The two-level random forest model is developed by incorporating error compensation. As shown in Fig. 3, the first-level random forest model  $f_{RF}^1()$  is first used to directly model the relationships between the population density and the covariables at the street scale. The expression can be written as

$$Pop(s) = f_{RF}^1(X(s)) + \epsilon'(s), \tag{1}$$

is the actual population density at street  $s$ ,  $X(s)$  is a vector representing the covariable values, and  $\epsilon'(s)$  is the error between the actual population density and the predicted values from the first-level random model.

Because of the existence of spatial dependency, the error from the first-level random forest model may still be related to the covariables. Therefore, the second random forest model  $f_{RF}^2()$  is further selected to model the relationship between the error term and the covariables. This error can be expressed as

$$\epsilon'(s) = f_{RF}^2(X(s)) + \epsilon''(s), \tag{2}$$

where  $\epsilon''(s)$  is the error between the error  $\epsilon'(s)$  from Eq. (1) and its corresponding predicted value from  $f_{RF}^2()$  at street  $s$ . That is, the second random forest model aims to predict the error term from the first random forest model. The two-level random forest model with the integration of the error compensation can enhance the prediction accuracy by making full use of the characteristics of spatial data. A detailed introduction to the two-level random forest model can be found in the related literature (Yang et al., 2023; Zheng et al., 2016).

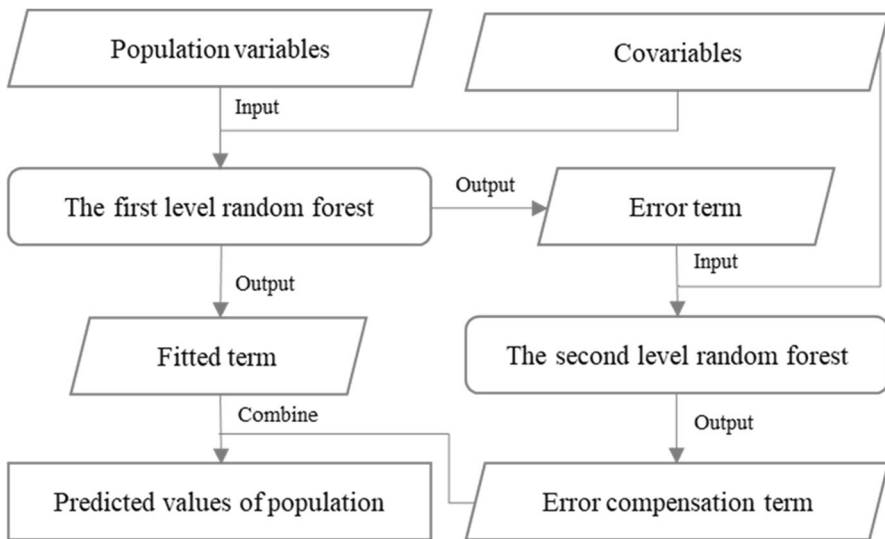


Fig. 3 Flowchart of the two-level random forest model



In contrast, the spatial distribution of the rural population is relatively simple, and the population is usually positively correlated with the number of buildings, whose functions are mainly used for living. Therefore, it is reasonable to estimate the population density in rural areas based on the areas of the buildings. High-resolution remote sensing images have become a useful tool for detecting land use/land cover due to their low cost, synoptic view, and repetitive data acquisition (Wei & Yang, 2020). Extracting building objects from high-resolution remote sensing images is not a difficult task, especially with the rapid development of deep learning technology. Among many deep learning algorithms, the U-Net convolutional neural network, with the advantages of fewer training samples and higher segmentation accuracy, are chosen in this research to obtain the spatial distribution of building image objects (Kim et al., 2019).

The U-Net convolutional neural network, which are achieved by extending a fully convolutional neural network, consist of two parts: an encoder and a decoder (Ronneberger et al., 2015). The encoder consists of four identical encoding blocks. Each block contains two convolutional layers with 3×3 kernels and a maximum pooling layer with 2×2 kernels (downsampling). The activation function is ReLU, the feature map size is halved, and the number of channels is doubled after each downsampling operation. The decoder also consists of four identical decoder blocks. Each block contains two convolutional layers with 3×3 kernels and a deconvolutional layer with 2×2 kernels (upsampling), and the result of the deconvolution is stitched with the feature map from the corresponding downsampling, followed by two convolutional operations with 3×3 kernels. The activation function is the same as that of ReLU, with the size of the feature map doubled and the number of channels halved after each upsampling operation. In the last layer of upsampling, a convolutional layer with a 1×1 kernel is used to map the feature vectors to the output layer of the network. The structure of the U-Net convolutional neural network is shown in Fig. 4 (Pan et al., 2020).

Therefore, the population number  $N(l)$  of spatial unit  $l$  in street unit  $r$  of rural areas can be calculated as

$$N(l) = area(l) \times \frac{Pop(r)}{area(r)}, \tag{3}$$

where  $Pop(r)$  is the total number of street unit  $r$  containing the spatial unit  $l$  that is recorded in the census, and  $area(r)$  and  $area(l)$  are the total areas of the building

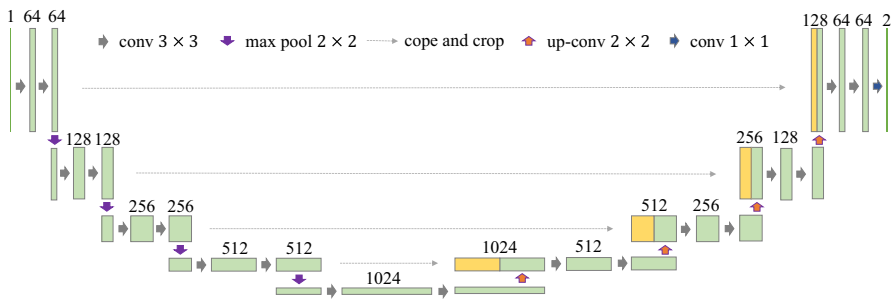


Fig. 4 Structure of the U-Net convolutional neural network

image objects of the street unit  $r$  and spatial unit  $l$ , respectively. The population numbers of different age groups in each grid can be estimated by the proportion of different age groups in each street unit.

Cross-validation was also used to evaluate the prediction results. Assuming that the populations of two different street units  $l_1$  and  $l_2$  are  $Pop(r_1)$  and  $Pop(r_2)$ , respectively, the predicted values of  $Pop(r_2)$  can be written as:

$$\widehat{Pop}(r_2) = area(r_2) \times \frac{Pop(r_1)}{area(r_1)}, \quad (4)$$

where  $area(r_1)$  and  $area(r_2)$  are the total areas of the building image objects. Based on the known value  $Pop(r_2)$  and the predicted value  $\widehat{Pop}(r_2)$ , the prediction accuracy for predicting the rural population can be obtained.

### A Distance-based Indicator Considering Heterogeneous Travel Modes for Measuring Spatial Accessibility

As mentioned above, spatial accessibility refers to the ability of residents to reach COVID-19 vaccination sites. Because different age groups have heterogeneous travel modes, a reasonable indicator to measure spatial accessibility should account for the comprehensive ability of different age groups to reach COVID-19 vaccination sites. Consequently, a hybrid time distance considering heterogeneous travel modes is presented to measure the spatial accessibility of each age group, and a population-weighted proximity indicator is then developed to evaluate the overall accessibility.

Travel modes mainly depend on the travel distance or time between the origin and destination. The greater the distance between the origin and destination is, the more likely the driving mode is to be chosen. The closer the distance is, the more likely the walking mode is to be chosen. Assuming that a distance threshold  $d$  for each individual is defined as the parameter for choosing the travel mode to COVID-19 vaccination sites, that is, when the actual distance is less than the distance threshold, an individual tends to choose the walking mode; otherwise, they choose the driving mode (only two primary travel modes, driving and walking, were considered in this study). Obviously, the distance threshold is a random variable, which is related to an individual's age, sex, occupation, etc. In this study, it is assumed that the distance threshold only varies across different age groups (e.g., children, middle-aged people, and older people) and follows a Gaussian distribution for each age group. Therefore, the conditional probability density  $g(d|i)$  of the distance threshold  $d$  for age group  $i$  ( $i = 1, 2$ , and  $3$  represent children, middle-aged people, and older people, respectively) can be expressed as

$$g(d|i) = \frac{1}{\sigma(i)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{d-\mu(i)}{\sigma(i)}\right)^2}, \quad (5)$$

where  $\mu(i)$  and  $\sigma(i)$  represent the mean and standard deviation of the distance threshold for each group, respectively. A survey about travel mode selection was used to validate the distribution functions and to estimate the parameters in this research.

For an individual in age group  $i$ , assume that the probability  $p(l|i)$  for selecting the driving mode is dependent on the network distance  $l$ ; then, a hybrid time distance  $HTD(l|i)$  ( $i=1, 2$ , and  $3$ ) with the integration of heterogeneous travel modes can be defined as:

$$HTD(l|i) = p(l|i)\frac{l}{u(i)} + (1 - p(l|i))\frac{l}{v(i)}, \quad (6)$$

where  $u(i)$  and  $v(i)$  represent the mean speeds of the driving and walking modes for different age groups, respectively. By considering that Eq. (5) essentially describes the conditional probability of travel mode selection,  $p(l|i)$  can be estimated as follows:

$$p(l|i) = \int_{-\infty}^l \frac{1}{\sigma(i)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu(i)}{\sigma(i)}\right)^2} dx. \quad (7)$$

The hybrid time distance only describes the ability of each group to reach COVID-19 vaccination sites. To describe the overall accessibility of residents from a certain origin, the population number of different groups should be combined with the hybrid time distance. Therefore, based on the hybrid time distance, a population-weighted proximity indicator is further proposed to measure the overall accessibility of each demand point to COVID vaccination sites. The population-weighted proximity indicator  $PWP(loc)$  at location  $loc$  can be defined as

$$PWP(loc) = \sum_{i=1}^3 \left( \frac{Pop(loc, i)}{\sum_{i=1}^3 Pop(loc, i)} \times HTD(d(loc)|i) \right), \quad (8)$$

where  $Pop(loc, i)$  is the total number of group  $i$ . Usually,  $d(loc)$  represents the nearest network distance between spatial locations  $loc$  and all the COVID-19 vaccination sites. It is worth noting that for a unit without road coverage, its nearest unit with road coverage is first searched based on the Euclidean distance, and the network distance between the unit and the vaccination site is then calculated by the sum of the Euclidean distance from the unit to its nearest unit and the network distance from the nearest unit to the vaccination sites. Compared with the hybrid time distance  $HTD(d, i)$ ,  $PWP(loc)$  can represent the impact of the age group on spatial accessibility.

## Results

### The Spatial Distribution of Population at a Fine Scale

First, based on the census statistical yearbook, the study area was divided into urban and rural areas at the street scale. For the spatial distribution of the urban population, the two-level random forest model was used to directly model the relationships between the population density and the covariables at the street scale. The covariables include the five land categories, the average nighttime light intensity, the average slope, the numbers of the five POI categories, and the road network density. Considering that the sample size of urban streets is only 27, which is too small and is difficult to effectively support random forest modeling at the street scale, all samples from urban and rural streets were selected to construct the two-level random forest model. Hence, the sample size is 70, which is the same as the number of street units. Tenfold cross-validation was applied to evaluate the performance of the prediction model. The sizes of the training and testing data are 63 and 7, respectively. To validate the effectiveness of the two-level random forest model, artificial neural networks (ANNs), support vector regression (SVR), and their corresponding two-level ANNs (TL-ANNs) and two-level SVR (TL-SVRs) were selected to conduct comparative analysis (Deng et al., 2017; Yang et al., 2018).

The root mean square error (RMSE) and the coefficient of determination ( $R^2$ ) were selected to evaluate the prediction accuracy. The statistical results are shown in Table 2. For one-level machine learning models, the  $R^2$  of 0.83 obtained using RF is larger than those of the ANNs and SVRs, which are 0.80 and 0.77, respectively. It can also be found that the  $R^2$  values obtained from two-level machine learning models (TL-ANNs, TL-SVR, TL-RF) are 0.83, 0.82, and 0.87, respectively, which are larger than those achieved using the one-level machine learning models. More importantly, the  $R^2$  of TL-RF is the largest among all the comparison models. Similar evaluation results can be found by using the RMSE. The RMSEs for TL-ANNs, TL-SVR, and TL-RF are 14,663, 15,339, and 12,901, respectively, and these values are smaller than those of their corresponding one-level models. Among all the models, the minimum error was achieved by the TL-RF model. The above results indicate that the TL-RF model can accurately model the relationships between the population density and the covariables at the

**Table 2** Comparison of the prediction accuracy of different methods

Method	$R^2$	RMSE
ANN	0.80	16,260
TL-ANN	0.83	14,663
SVR	0.77	18,308
TL-SVR	0.82	15,339
RF	0.83	15,411
TL-RF	0.87	12,901

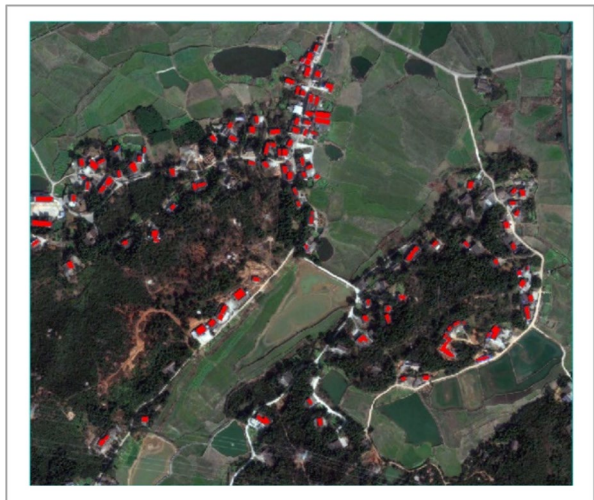
street scale. The built TL-RF model was then used to predict the spatial distribution of the urban population of each grid unit with the corresponding covariables.

For the spatial distribution of the rural population, the U-Net convolutional neural networks were used to extract the building image objects. This process was directly implemented in the software ENVI 5.6. The ROIs (regions of interest) tool was first used to generate building image samples or objects. As shown in Fig. 5, the building image object was labelled as a polygon in red. The total number of samples is 3,000, and the training samples and test samples account for 80% and 20% of the total number of samples, respectively. The kappa coefficient was chosen to evaluate the classification accuracy. The value of the kappa coefficient was 0.93, which indicated that the U-Net convolutional neural networks could effectively extract the building image objects. The final result of building image objects extracted from the whole rural area is shown in Fig. 6. In addition, the areas of building image objects and the population number were calculated for each grid unit. According to Eq. (4), the  $R^2$  and RMSE of the cross-validation results are 0.80 and 9,035, respectively, indicating that this method is effective in estimating the rural population. According to the age structure of each age group, the numbers of each age group were calculated. Figure 7 shows the spatial distribution of the total population and three age groups in the whole study area.

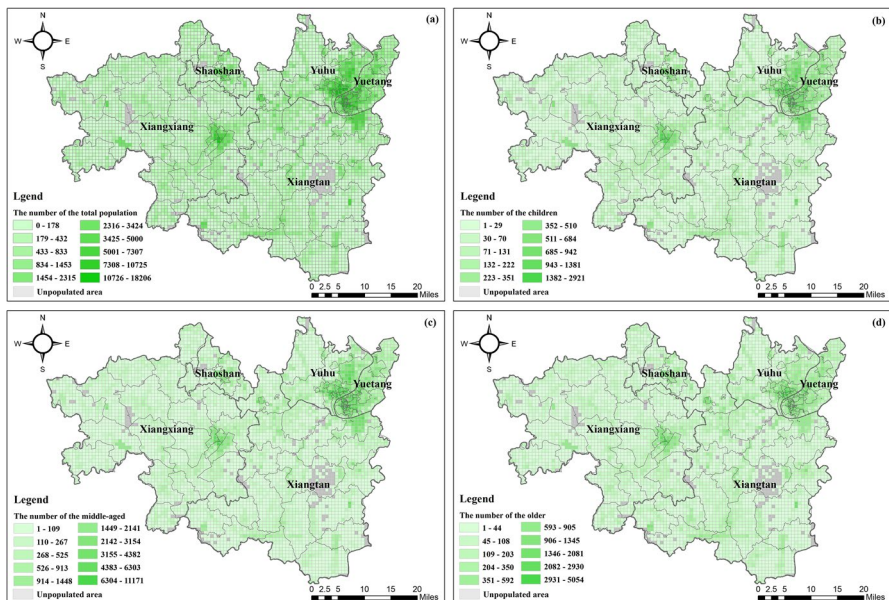
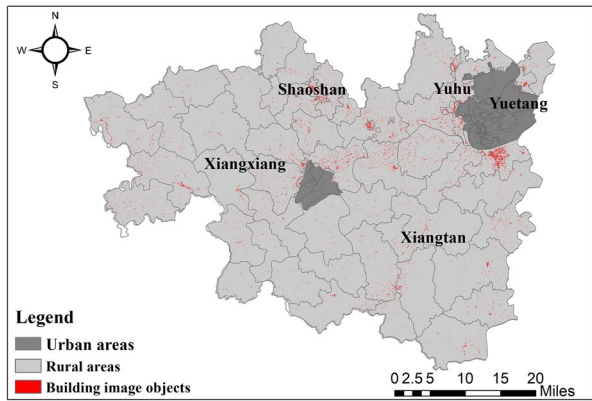
### The Evaluation of Spatial Accessibility

For each age group, 100 samples were collected to obtain the choice of the travel modes. The histogram and statistical results of the distance threshold for choosing different travel modes are shown in Fig. 8. The mean distance threshold for the children is 14.4 min, followed by the older people (21.3 min) and the middle-aged people (29.4 min). Meanwhile, the standard deviations shows the same order. The minimum value obtained for the children is 6.5 min, followed by the older people

**Fig. 5** Building image samples are in red



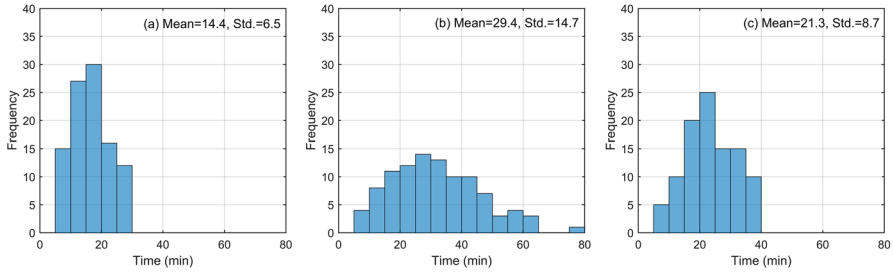
**Fig. 6** Spatial distribution of building image objects



**Fig. 7** Fine-scale population distributions of different groups: **a** the total population, **b** the children, **c** the middle-aged people, and **d** the older people

(8.7 min) and the middle-aged people (14.7 min). This indicates that the children tend to be driven and that the middle-aged people can tolerate long walks in the study area.

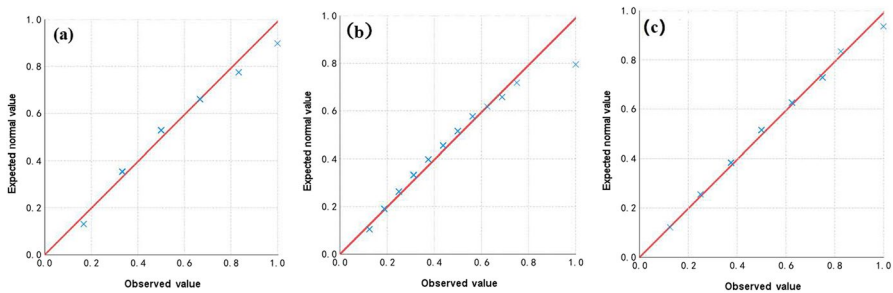
Quantile–quantile plots (QQ plots) were used to verify whether the distance threshold distributions of different groups followed a Gaussian distribution. As



**Fig. 8** Histograms of the distance thresholds for selecting different travel modes for different groups: **a** The children; **b** the middle-aged people; and **c** the older people

shown in Fig. 9, the scatter plots in the three subfigures are approximately distributed near the  $y=x$  line, which means that the three groups of samples approximately follow a normal distribution. Based on Eq. (5), the conditional probability density of the distance threshold for each group can be estimated. Next, according to the walking trajectory data, the mean speeds by driving or walking modes for different age groups were calculated. As shown in Table 3, the mean speeds of walking  $u(i)$  for the children, middle-aged people, and older people are 2.74, 3.89, and 3.28 km/h, respectively. The mean speed of driving modes  $v(i)$  for all groups is the same (40 km/h), which was obtained by the speed limit on the regional roads.

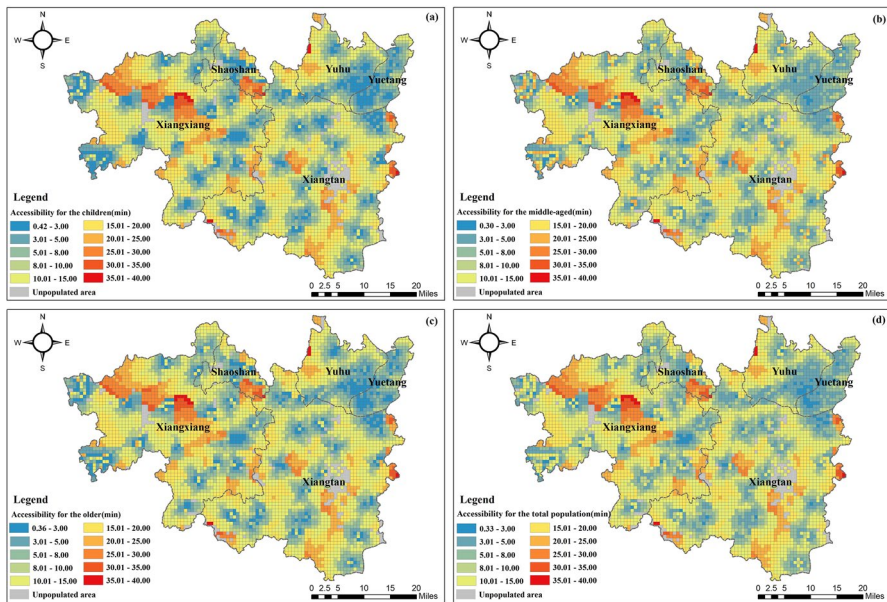
Based on the hybrid time distance and the population-weighted proximity indicator, the final results of spatial accessibility were calculated for each age group and the total population at each grid, respectively. Figure 10 shows the spatial distribution of accessibility. From Figs. 10(a) to (c), it can be found that the spatial distribution of accessibility for each group presents almost the same clustering patterns. Although the travel modes of these three groups are inconsistent, this



**Fig. 9** QQ plots of the distance thresholds for different groups: **a** The children; **b** the middle-aged people; and **c** the older people

**Table 3** Characteristics of travel mode for different age groups

Mean speed of different travel modes	Group		
	Children	Middle-aged	Older
$u(i)$ (km/h)	2.74	3.89	3.28
$v(i)$ (km/h)	40	40	40



**Fig. 10** Spatial distributions of accessibility for different groups: **a** children, **b** middle-aged, **c** older, and **(d)** the total population

difference does not obviously change the spatial distribution of accessibility. The areas with low values of spatial accessibility are mainly located around the vaccination sites, while the areas with high values of spatial accessibility are located far away from the vaccination sites. This pattern is also shown in Fig. 10(d), which describes the weighted accessibility for the total population. Because of the same pattern of spatial accessibility for each group, the population-weighted results based on Eq. (8) also show the same pattern.

The statistical results of accessibility for each group travel mode are shown in Table 4. For the whole study area, the mean and standard deviation of spatial accessibility are 12.09 and 6.65 min, respectively, which are almost the same as the values for each group. Overall, there is no significant difference in spatial accessibility for different groups in the whole area study. However, according to regional statistical data, spatial heterogeneity for the accessibility of each group and the total population is shown. As shown in Table 4, for the children, Xiangxiang County has the largest spatial accessibility (13.05 min), followed by Shaoshan County (12.66 min), Xiangtan County (12.17 min), Yuhu District (10.27 min), and Yuetang District (6.07 min). For the middle-aged and the older groups, the order of the mean values of spatial accessibility is similar to that for the children. The area with the largest mean value is Xiangxiang County, and that with the lowest mean value is Yuetang District. According to the population-weighted result, Xiangxiang County and Yuetang District correspond to the maximum and minimum values of the overall accessibility, respectively. This shows that residents in



**Table 4** Statistics of spatial accessibility to COVID-19 vaccination sites

Area	Number of streets	Number of grids	Group							
			Children		Middle-aged		Older		Total population	
			Mean (min)	Std (min)	Mean (min)	Std (min)	Mean (min)	Std (min)	Mean (min)	Std (min)
All	70	5,068	11.67	6.97	12.31	6.73	11.89	6.86	12.09	6.65
Shaoshan	4	215	12.01	8.34	12.96	7.87	12.37	8.12	12.66	7.85
Xiangtan	17	2,093	11.92	6.27	12.32	6.05	12.01	6.21	12.17	6.03
Xiangxiang	22	2,052	12.44	7.58	13.34	7.27	12.79	7.40	13.05	7.12
Yuetang	12	233	5.68	3.09	6.28	2.74	5.88	2.94	6.07	2.76
Yuhu	15	475	10.01	6.22	10.42	5.99	10.10	6.13	10.27	6.00

Yuetang County more easily reach vaccination sites, and residents in Xiangxiang County generally need more time to access vaccination sites.

Meanwhile, the standard deviation represents the variability in an area and describes the spatial fairness of vaccination sites. For the spatial accessibility of the children, the maximum standard deviation is 8.34 min in Shaoshan County, the minimum standard deviation is 3.09 min in Yuetang District, and the other three areas have similar standard deviations. The standard deviations of spatial accessibility of the other two groups have similar patterns. Similar results can also be found in terms of the accessibility for the total population. The corresponding standard deviations of Shaoshan County and Yuetang District are the largest and smallest, with values of 7.85 and 2.76 min, respectively. The standard deviations of Xiangxiang County (7.12 min), Xiangtan County (6.03 min), and Yuhu District (6.00 min) are considered intermediate, and the standard deviation of the overall accessibility in Xiangxiang County is slightly larger than that of the other two counties. This means that Shaoshan County is the most unbalanced area of accessibility, and Yuetang District is the most balanced area.

## Discussion

In this study, a new framework based on fine-scale population distributions and heterogeneous travel modes is proposed to evaluate the spatial distribution of accessibility to COVID-19 vaccination sites. Two critical tasks are involved in this framework: one is to obtain the population distribution at a fine scale, and the other is to construct a population-weighted proximity indicator with the integration of heterogeneous travel modes of different age groups. Xiangtan city, China, is selected as the study area. The major findings can be summarized as follows.

First, the divide-and-conquer strategy presented in this research has been proven to be effective in estimating the urban and rural population. Specifically, the two-level random forest model is used to predict the urban population, and the  $R^2$  and RMSE from this model are 0.87 and 12,901, respectively. The  $R^2$  (RMSE) from the two-level random forest model is larger (smaller) than those from the comparison models, which demonstrates that the two-level random forest model outperforms the comparison models, and a more accurate distribution of the urban population has been obtained. Considering the positive correlation between population and building area in rural areas, U-Net convolutional neural networks are used to extract the building objects from high-resolution remote sensing images, and the fine-scale population in the rural area is predicted according to the building area. The  $R^2$  and RMSE are 0.80 and 9,035, respectively, which indicates that U-Net convolutional neural networks can indirectly obtain a more accurate population distribution in rural areas.

Second, the spatial distribution of accessibility to COVID-19 vaccination sites for different age groups exhibits a similar pattern. Travel modes are considered to be a critical factor in evaluating spatial accessibility. In this study, it is assumed that travel modes vary across different age groups. Based on a survey about travel mode choice, walking trajectory, and road network, the probability and characteristics of

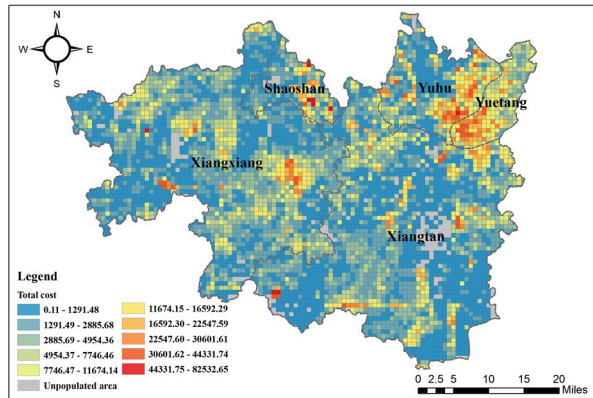
travel modes of different age groups are estimated, and a hybrid time distance is then defined to measure spatial accessibility for each age group. Figure 4 shows that the spatial distribution of the hybrid time distance is similar in different age groups. The areas with low values are mainly located around the vaccination sites, and those with high values are far from the vaccination sites. As shown in Fig. 10 and Table 4, the probability and characteristics of travel modes vary with different age groups, but this difference does not significantly affect the spatial distribution of accessibility to COVID-19 vaccination sites, and only some local areas show inconsistency.

Third, at the county level, inequity in spatial accessibility to COVID-19 vaccination sites has been confirmed in the study area. The hybrid time distance is the only measure of spatial accessibility for each age group. Therefore, this study further proposes the population-weighted proximate indicator to evaluate the overall accessibility at a certain grid. By aggregating the grid into county scales, the inequity in the overall accessibility is reflected in two levels: intercounty and intracounty. At the intercounty level, it generally takes less time for residents in Yutang District to reach COVID-19 vaccination sites, while it usually takes longer for those in Xiangxiang County. At the intracounty level, Shaoshan County has the largest standard deviation of overall accessibility, so it faces serious inequities. The Yuetang District has the smallest variation in overall accessibility, which means relative equities in spatial accessibility to COVID-19 vaccination sites.

The spatial allocation of medical resources is a significant component of the prevention and control of the COVID-19 epidemic. The ultimate objective for evaluating spatial accessibility is to provide public health policy-makers with fundamental decision information on the spatial optimization of vaccination sites. According to the spatial distribution of accessibility to COVID-19 vaccination sites, the inequity in spatial accessibility to COVID-19 vaccination sites in the study area has been confirmed. Two special areas, namely, Yuetang District and Xiangxiang County, are identified as having the maximum and minimum accessibility, respectively, and the variations in accessibility in Yuetang District and Shaoshan County also correspond to the minimum and maximum values, respectively. Thus, Shaoshan County should pay more attention to its relatively low mean and highest variance of spatial accessibility. Improving regional transportation or increasing vaccination sites may be key strategies to reduce accessibility inequities.

The population-weighted proximity indicator used to measure spatial accessibility is only applicable to the travel time distance from the resident's location (a grid unit) to the nearest vaccination sites. Although population distribution information has been considered in the proposed indicator, the demand capacity, namely, the total population number, is neglected. To further understand the time cost of the resident, spatial accessibility with the consideration of the population number or the demand capacity is computed. The spatial distribution of the total cost at each grid is shown in Fig. 11. The areas with high values of the total cost exhibit obvious characteristics of the aggregation distribution, which are mainly distributed in the urban areas. Rural areas generally correspond to a relatively low total cost. The spatial distribution of the total cost is similar to that of the total population shown in Fig. 7. Because the total cost comes from the product of spatial accessibility and the population in each grid and the magnitude of the population is much higher than

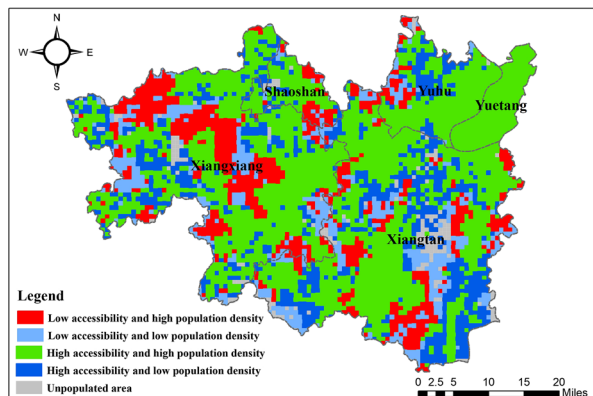
**Fig. 11** Spatial distribution of the total cost



that of spatial accessibility, the final total cost is mainly dependent on the total population. Hence, the population distribution should be the most important factor in determining the spatial locations of vaccination sites.

On the basis of the “15-min city” (Pinto & Akhavan, 2022) and the threshold population density classification (Fujibe, 2012), the grid is divided into four categories: low accessibility ( $> 15$  min) and low population density ( $\leq 100$  per square kilometre), low accessibility ( $> 15$  min) and high population density ( $> 100$  per square kilometre), high accessibility ( $\leq 15$  min) and low population density ( $\leq 100$  per square kilometre), and high accessibility ( $\leq 15$  min) and high population density ( $> 100$  per square kilometre). The classification results are shown in Fig. 12. The spatial distributions of these four patterns also show aggregation characteristics. The areas with high accessibility and high population density are the largest, those with low accessibility and low population density are the lowest, and the areas of the other two types are between them. Actually, the areas with low accessibility and high population density marked in red colour in Fig. 12 need to be areas of focus, as it takes more time for a large population to reach vaccination sites. These red areas

**Fig. 12** Classification of accessibility and population



are mainly located in Xiangxiang and Xiangtan counties. For these areas, effective measures should be taken to improve spatial accessibility to vaccination sites. For example, temporary vaccination sites can be added in these areas.

Several limitations need to be pointed out. First, this study did not deeply discuss the prediction of population for different age groups at the grid-unit level. On the basis of the fine-scale population distribution, the numbers of different age groups in each grid were estimated from the proportions of different age groups in the street units containing the grid, which have been recorded in the demographic yearbook. Therefore, it is assumed that the age structure of the population at each grid in each street is consistent. Due to the lack of age-related data at fine spatial scales in this research, the above assumption is difficult to test. Therefore, how to estimate the age structure at a fine scale is an important part of future work.

Second, the inaccurate road network from OpenStreetMap may lead to inaccurate spatial accessibility results. As one of the most successful products of volunteered geographic information, OpenStreetMap is mainly contributed by volunteers, most of whom are not professionals or are untrained (Muttaqien et al., 2018). Therefore, the spatial data quality relevant to the attribute and positional accuracy of OSM data is a critical factor that influences the results of spatial accessibility, which should be further considered in the future.

Third, due to the limitation of data, this study only considers two types of travel modes and simplifies the complex travel modes. At present, bike-sharing is a very important travel mode in cities that was not included in the travel time analysis. Because of the diversity of driving modes, such as taxis, buses, and cars, how these different modes impact the travel time from the resident's location to the vaccination sites is also unknown. Additionally, travel mode selection is only related to the travel time and is not involved in cost or group preference. It may be more reasonable to combine different factors to determine the selection of travel modes, which should be discussed in future work.

## Conclusions

The purpose of this study is to evaluate spatial accessibility to COVID-19 vaccination sites based on fine-scale population distributions and heterogeneous travel modes. To achieve this objective, the divide-and-conquer strategy is first proposed to predict the fine-scale populations of urban and rural areas. This strategy for predicting the spatial distribution of population can enrich the theory and methods of spatial analysis to address spatial heterogeneity. With the integration of heterogeneous travel modes of different age groups, a hybrid time distance and a population-weighted proximity indicator are then defined to measure spatial accessibility to COVID-19 vaccination sites for each age group and the total population. The proposed hybrid time and related indicator can be transferred to relevant studies such as time geography or human behaviour patterns. Meanwhile, because population distributions or transportation patterns are key elements in evaluating spatial accessibility, this framework can be used not only to evaluate the accessibility of vaccination

sites but also to analyse the spatial accessibility of other facilities or services, such as commercial and sports facilities.

Xiangtan city, China, was chosen as the study area. It can be found that by using the divide-and-conquer strategy, a relatively high accuracy of the population dataset at a fine scale could be generated for the study area. The fine-scale population dataset can support relevant research, including urban planning and commercial site selection. On the basis of the spatial distribution of the population, the result of spatial accessibility indicates that the spatial distribution of accessibility to COVID-19 vaccination sites for different age groups exhibits a similar pattern, and inequities in spatial accessibility to COVID-19 vaccination sites have been confirmed. In particular, areas with low accessibility and high population density were identified, which can provide policy-makers and planners of public health with important information on optimizing the spatial patterns of existing COVID-19 vaccination sites.

**Acknowledgements** This study was supported by the Science and Technology Innovation Program of Hunan Province (No. 2023sk2081).

**Data Availability** Data will be made available on request.

## Declarations

**Conflicts of Interest** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, 395(10228), 931–934. [https://doi.org/10.1016/S0140-6736\(20\)30567-5](https://doi.org/10.1016/S0140-6736(20)30567-5)
- Bakillah, M., Liang, S., Mobasheri, A., Arsanjani, J. J., & Zipf, A. (2014). Fine-resolution population mapping using OpenStreetMap points-of-interest. *International Journal of Geographical Information Science*, 28(9), 1940–1963. <https://doi.org/10.1080/13658816.2014.909045>
- Bucyibaruta, G., Blangiardo, M., & Konstantinoudis, G. (2022). Community-level characteristics of COVID-19 vaccine hesitancy in England: A nationwide cross-sectional study. *European Journal of Epidemiology*, 37(10), 1071–1081. <https://doi.org/10.1007/s10654-022-00905-1>
- Deng, M., Yang, W., Liu, Q., & Zhang, Y. (2017). A divide-and-conquer method for space-time series prediction. *Journal of Geographical Systems*, 19(1), 1–19. <https://doi.org/10.1007/s10109-016-0241-y>
- Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolessi, L., Squires, J., Borkovska, O., & Baptista, S. R. (2015). Taking advantage of the improved availability of census data: A first look at the gridded population of the world, version 4. *Papers in Applied Geography*, 1(3), 226–234. <https://doi.org/10.1080/23754931.2015.1014272>

- Duhon, J., Bragazzi, N., & Kong, J. D. (2021). The impact of non-pharmaceutical interventions, demographic, social, and climatic factors on the initial growth rate of COVID-19: A cross-country study. *Science of the Total Environment*, 760, 144325. <https://doi.org/10.1016/j.scitotenv.2020.144325>
- Fujibe, F. (2012). Dependence of long-term temperature trends on wind and precipitation at urban stations in Japan. *Journal of the Meteorological Society of Japan. Ser. I*, 90(4), 525–534. <https://doi.org/10.2151/jmsj.2012-406>
- Guhlincozzi, A. R., & Lotfata, A. (2021). Travel distance to flu and COVID-19 vaccination sites for people with disabilities and age 65 and older, Chicago metropolitan area. *Journal of Research in Health Sciences*. <https://doi.org/10.1108/JHR-03-2021-0196>
- Jia, P., & Gaughan, A. E. (2016). Dasymetric modeling: A hybrid approach using land cover and tax parcel data for mapping population in Alachua County, Florida. *Applied Geography*, 66, 100–108. <https://doi.org/10.1016/j.apgeog.2015.11.006>
- Kaufman, B. G., Whitaker, R., Mahendraratnam, N., Hurewitz, S., Yi, J., Smith, V. A., & McClellan, M. (2021). State variation in effects of state social distancing policies on COVID-19 cases. *BMC Public Health*, 21, 1239. <https://doi.org/10.1186/s12889-021-11236-3>
- Kim, J. H., Lee, H., Hong, S. J., Kim, S., Park, J., Hwang, J. Y., & Choi, J. P. (2019). Objects segmentation from high-resolution aerial images using U-net with pyramid pooling layers. *IEEE Geoscience & Remote Sensing Letters*, 16(1), 115–119. <https://doi.org/10.1109/LGRS.2018.2868880>
- Kwan, M., Murray, A. T., O’Kelly, M. E., & Tiefelsdorf, M. (2003). Recent advances in accessibility research: Representation, methodology and applications. *Journal of Geographical Systems*, 5(1), 129–138. <https://doi.org/10.1007/s101090300107>
- Li, C. K., Pu, P., Fang, J., Ma, K. S., & Chen, A. M. (2018). Research on accessibility evaluation of medical service in Xiangtan city based on improved gravity model. *Economic Geography*, 38(12), 83–88. <https://doi.org/10.15957/j.cnki.jjdl.2018.12.011>
- Liu, D., Kwan, M. P., Kan, Z., Song, Y., & Li, X. (2022). Racial/Ethnic inequity in transit-based spatial accessibility to COVID-19 vaccination sites. *Journal of Racial and Ethnic Health Disparities*. <https://doi.org/10.1007/s40615-022-01339-x>
- Liu, H., Xu, Y., Tang, J., Deng, M., Huang, J., Yang, W., & Wu, F. (2020). Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities. *Transactions in GIS*, 24(5), 1359–1381. <https://doi.org/10.1111/tgis.12642>
- Liu, Z., Ma, T., Du, Y., Pei, T., Yi, J., & Peng, H. (2018). Mapping hourly dynamics of urban population using trajectories reconstructed from mobile phone records. *Transactions in GIS*, 22(2), 494–513. <https://doi.org/10.1111/tgis.12323>
- Muttaqien, B. I., Ostermann, F. O., & Lemmens, R. L. G. (2018). Modeling aggregated expertise of user contributions to assess the credibility of OpenStreetMap features. *Transactions in GIS*, 22, 823–841. <https://doi.org/10.1111/tgis.12454>
- Pan, G., Zheng, Y., Guo, S., & Lv, Y. (2020). Automatic sewer pipe defect semantic segmentation based on improved U-net. *Automation in Construction*, 119, 103383. <https://doi.org/10.1016/j.autcon.2020.103383>
- Pinto, F., & Akhavan, M. (2022). Scenarios for a post-pandemic city: Urban planning strategies and challenges of making “Milan 15-minutes city.” *Transportation Research Procedia*, 60, 370–377. <https://doi.org/10.1016/j.trpro.2021.12.048>
- Qi, F., Barragan, D., Rodriguez, M. G., & Lu, J. (2022). Evaluating spatial accessibility to COVID-19 vaccine resources in diversely populated counties in the United States. *Frontiers in Public Health*, 10, 895538. <https://doi.org/10.3389/fpubh.2022.895538>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *International conference on medical image computing and computer-assisted intervention* (pp. 234–241). Springer.
- Tao, R., Downs, J., Beckie, T. M., Chen, Y., & McNelley, W. (2020a). Examining spatial accessibility to COVID-19 testing sites in Florida. *Annals of GIS*, 26(4), 319–327. <https://doi.org/10.1080/19475683.2020.1833365>
- Tao, Z., Cheng, Y., Du, S., Feng, L., & Wang, S. (2020b). Accessibility to delivery care in Hubei Province, China. *Social Science & Medicine*, 260, 112406. <https://doi.org/10.1016/j.socscimed.2020.113186>
- Wei, D., & Yang, W. (2020). Detecting damaged buildings using a texture feature contribution index from post-earthquake remote sensing images. *Remote Sensing Letters*, 11(2), 127–136. <https://doi.org/10.1080/2150704X.2019.1692389>
- Xiao, T., Ding, T., Zhang, X., Tao, Z., & Liu, Y. (2022). Spatial accessibility to sports facilities in Dongguan, China: A multi-preference gaussian two-step floating catchment area method. *Applied Spatial Analysis and Policy*, 15(4), 1093–1114. <https://doi.org/10.1007/s12061-022-09436-4>

- Yang, W., Deng, M., Xu, F., & Wang, H. (2018). Prediction of hourly PM<sub>2.5</sub> using a space-time support vector regression model. *Atmospheric Environment*, *181*, 12–19. <https://doi.org/10.1016/j.atmosenv.2018.03.015>
- Yang, W., Wan, X., Liu, M., Zheng, D., & Liu, H. (2023). A two-level random forest method for predicting the population distributions of urban functional zones: A case study in Changsha, China. *Sustainable Cities and Society*, *88*, 104297. <https://doi.org/10.1016/j.scs.2022.104297>
- Yao, Y., Liu, X., Li, X., Zhang, J., Liang, Z., Mai, K., & Zhang, Y. (2017). Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. *International Journal of Geographical Information Science*, *31*(6), 1220–1244. <https://doi.org/10.1080/13658816.2017.1290252>
- Zeng, C., Zhou, Y., Wang, S., Yan, F., & Zhao, Q. (2011). Population spatialization in China based on night-time imagery and land use data. *International Journal of Remote Sensing*, *32*(24), 9599–9620. <https://doi.org/10.1080/01431161.2011.569581>
- Zheng, D., Hu, W., & Li, P. (2016). Predicting ionospheric critical frequency of the F2 layer over Lyckesele using the neural network improved by error compensation technology. *Survey Review*, *48*(347), 130–139. <https://doi.org/10.1179/1752270615Y.0000000015>
- Zhou, Y., Ma, M., Shi, K., & Peng, Z. (2020). Estimating and interpreting fine-scale gridded population using random forest regression and multisource data. *ISPRS International Journal of Geo-Information*, *9*(6), 369. <https://doi.org/10.3390/ijgi9060369>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Wentao Yang<sup>1,2</sup> · Fengjie Wang<sup>1,2</sup> · Yihan You<sup>2</sup> · Xiafan Wan<sup>1,2</sup> · Sijie Cheng<sup>3</sup> · Zhixiong Fang<sup>4,5</sup>

✉ Zhixiong Fang  
fzx8214907@sina.com

Wentao Yang  
yangwentao8868@126.com

Fengjie Wang  
1945234876@qq.com

Yihan You  
2300576285@qq.com

Xiafan Wan  
1023292673@qq.com

Sijie Cheng  
749896315@qq.com

<sup>1</sup> National-Local Joint Engineering Laboratory of Geospatial Information Technology, Hunan University of Science and Technology, Xiangtan 411201, China

<sup>2</sup> School of Earth Science and Spatial Information Engineering, Hunan University of Science and Technology, Xiangtan 411201, China

<sup>3</sup> Economics and Management School, Wuhan University, Wuhan 430072, China

<sup>4</sup> School of Physical Education, Hunan University of Science and Technology, Xiangtan 411201, China

<sup>5</sup> Xiangtan Central Hospital, Xiangtan 411199, China