

# A two-stage location model covering COVID-19 sampling, transport and DNA diagnosis: design of a national scheme for infection control

Wang Fei,<sup>1</sup> Lv jiamin,<sup>2</sup> Wang Chunting,<sup>3</sup> Li Yuling,<sup>4</sup> Xi Yueting<sup>5</sup>

<sup>1</sup>School of Economics and Management, Chang'an University, China; <sup>2</sup>School of Economics and Management, Chang'an University, China; <sup>3</sup>School of Economics and Management, Chang'an University, China; <sup>4</sup>School of Economics and Management, Chang'an University, China; <sup>5</sup>School of Economics and Management, Chang'an University, China

## Abstract

During the COVID-19 pandemic, a system was established in China that required testing of all residents for COVID-19. It consisted of sampling stations, laboratories capable of carrying out DNA investigations and vehicles carrying out immediate transfer of all samples from the former to the latter. Using Beilin District, Xi'an City, Shaanxi Province, China as example, we designed a genetic algorithm based on a two-stage location coverage model for the location of the sampling stations with regard to existing residencies as well as the transfer between the sampling stations

and the laboratories. The aim was to estimate the minimum transportation costs between these units. In the first stage, the model considered demands for testing in residential areas, with the objective of minimizing the costs related to travel between residencies and sampling stations. In the second stage, this approach was extended to cover the location of the laboratories doing the DNA-investigation, with the aim of minimizing the transportation costs between them and the sampling stations as well as the estimating the number of laboratories needed. Solutions were based on sampling stations and laboratories existing in 2022, with the results visualized by geographic information systems (GIS). The results show that the genetic algorithm designed in this paper had a better solution speed than the Gurobi algorithm. The convergence was better and the larger the network size, the more efficient the genetic algorithm solution time.

Correspondence: Xi Yueting, School of Economics and Management, Chang 'an University, Xi 'an, Shaanxi Province, China.

Tel.: +86.151.2903.8805

E-mail: 28220546@qq.com

Key words: COVID-19DNA testing, two-stage location model, genetic algorithm, China.

Conflict of interest: the authors declare no potential conflict of interest, and all authors confirm accuracy.

Funding: this article was supported by the Natural Science Foundation of China "Research on Urban Spatial Structure Reconstruction Based on Modern Logistics" (41371180).

Availability of data and materials: all data generated or analyzed during this study are included in this published article.

Received: 12 March 2024.

Accepted: 9 September 2024.

©Copyright: the Author(s), 2024

Licensee PAGEPress, Italy

Geospatial Health 2024; 19:1281

doi:10.4081/gh.2024.1281

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

*Publisher's note: all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.*

## Introduction

The coronavirus disease 2019 (COVID-19) pandemic spread rapidly worldwide after the first cases in December 2019 were detected in Wuhan, Hubei Province in central China (Zhou *et al.*, 2020). This highly contagious virus has had a major impact on individuals, societies and economies, causing infections and even deaths worldwide (Nicola *et al.*, 2020; Davis *et al.*, 2023). To mitigate casualties and economic loss, there is a pressing need to establish service facilities for urgent patient support, such as evacuating or relocating victims to various emergency medical care facilities, shelters, mobile hospitals, etc. Structural imbalances of the current healthcare system, e.g., limited medical service supply, uneven spatial distribution and underutilization of resources, pose challenges to the growing demand for essential services.

The high transmissibility of the coronavirus necessitates timely screening and rapid isolation of infected individuals propelling a strong demand for the location of testing services and emergency facilities. This requires the establishment of temporary sampling points in various districts followed by the transfer of samples to laboratories designated for DNA investigations. When the workload surged in the initial stages of the COVID-19 pandemic, some regions in China experienced suboptimal positioning of sampling stations hindering the timely identification and control of infection sources. Moreover, the efficiency of sample processing was often influenced by the distance between sampling stations and laboratories capable of investigating virus DNA. These issues reveal location deficiencies for which we propose a two-stage model covering the location of sampling stations and laboratories, with the aim to facilitate implementation of rapid and accurate testing followed by DNA investigation.



The facility location problem is one of the classical problems in combinatorial optimization. Because it is difficult to obtain the optimal solution in such situations, heuristic algorithms are often used. Compared to other algorithms, this particular approach has stronger universality and searching ability and is suitable for solving large-scale complex location problems. Genetic heuristic algorithms (GAs) are inspired by the principle of natural selection and used for solving both constrained and unconstrained optimization problems. GAs assist the solution of complex problems by improving a population of potential solutions iteratively. They have been widely used in emergency facility location (Murali *et al.*, 2016; Liu *et al.*, 2016), emergency logistics planning (Nikzami *et al.*, 2020; Wang *et al.*, 2021) and other fields. Although the GAs have shortcomings, they can be improved by technical design (Bashiri *et al.*, 2018; Wang *et al.*, 2022; Liu *et al.*, 2023) or through combination with other algorithms (Farrokhi-Asl *et al.*, 2017; Rabbani *et al.*, 2018; Zhong *et al.*, 2020).

## Literature review

Work on facility locations is typically categorized into different models, such as the discrete coverage approach (the P-median model) and efforts to minimize the maximal distance for all demand points (the P-centre model). Most facility location studies concentrate on the requirement of facility points to identify needs that cover the largest proportion. In the context of emergency medical service facilities, the objective function commonly includes maximizing the coverage level, minimizing the number of facilities and the transportation costs (Karimi *et al.*, 2011; Nogueira *et al.*, 2016; Alizadeh & Nishi, 2020).

Ahmadi-Javid *et al.* (2016) conducted a comprehensive study on the location of medical facilities considering ten dimensions that included uncertainty, periodicity, and discreteness leading to a location model with targeted action. Liu *et al.* (2016) focused on mobile hospitals, while Bélanger *et al.* (2020) emphasized ambulance deployment. Siddiq *et al.* (2013) studied the coverage of service radii of Automatic External cardiac Defibrillators (AEDs) in metropolitan Toronto, Canada and found that improved service radii significantly improved survival rates after out-of-hospital cardiac arrests. Schneeberger *et al.* (2016), based on previous work by Gendreau *et al.* (1997), constructed a single-period, dual-coverage model leading to re-siting of available resources, while Gu *et al.* (2018) solved a similar situation using a 'greedy algorithm' to find the preferable of temporary emergency medical service stations. Dong *et al.* (2018), on the other hand, utilized different accident emergency response times in Dalian, China as constraints in a study of emergency facility subway locations that resulted in a lower number of sites.

Caunhyea *et al.* (2016) proposed a two-stage path dealing with various uncertainty conditions and Mantzaras *et al.* (2017) constructed an optimization model with the objective of minimizing the cost of an infectious waste management network, which resulted in optimal locations of transfer stations and treatment centres. Pichka *et al.* (2018) used the minimum cost as the objective function, incorporating a third-party logistics provider to develop a variable two-stage open model, while Sancu *et al.* (2019) proposed a two-stage, stochastic planning model that not only considered facility locations, but also equipment location. Oksuz *et al.* (2020) developed a two-stage stochastic planning model with the objective of minimising the total TMC set-up cost and the expected total transport cost by considering the distance between the disaster area and the medical centres and assigning different casualty categories

to these medical centres for emergency medical response. The siting and transportation flow with regard to a healthcare temporary storage centre for waste was approached by Liu *et al.* (2020) with a two-stage model considering environmental traffic and vehicle loading factors. Amideo *et al.* (2021) studied the siting of shelters in different contexts by first focusing on location followed by a second stage optimizing the evacuation routes. Seraji *et al.* (2021) constructed a two-stage, multi-objective model for the study of the problem of location and uneven distribution of refuge places. Here, the first stage considered the various distances between warehouse and refuge, with the second suggesting a distribution that minimized the cost caused by unequal distribution systems, while Long *et al.* (2023) constructed a similar stochastic model considering pre-pandemic preparatory measures and post-pandemic rescue operations based on a preference-inspired co-evolutionary algorithm that was validated by an actual case in Hunan Province, China.

There is a limited number of studies on emergency location models utilizing complete coverage approaches and there is scarcity of work considering the capacity limitations of facility points and service radii. The potential contributions of this study include: i) development of a versatile, two-stage location model that takes capacity constraints and rapid response into account; ii) creation of a genetic algorithm that effectively helps solving this problem; and iii) validation of the effectiveness of the algorithm and model. The approach is summarised in a flow chart (Figure 1).

## Materials and Methods

To cover the medical needs for all regions, reduce the construction cost as much as possible and improve the efficiency of medical testing as well as the convenience of residents, it was necessary to optimise the sampling system in the entire region and scientifically select the number and location of sampling points. This was followed by the selection of DNA laboratory locations taking into account the integrity of medical testing according to Standard Operating Procedures (SOPs). We also considered the transport costs at this stage. These operations were eventually transformed into a mathematical model. In the first stage, the location of the network of sampling points was determined based on the location and number of residences. In the second stage, the location of the laboratories was decided based on the location of the sampling points.

## Study site

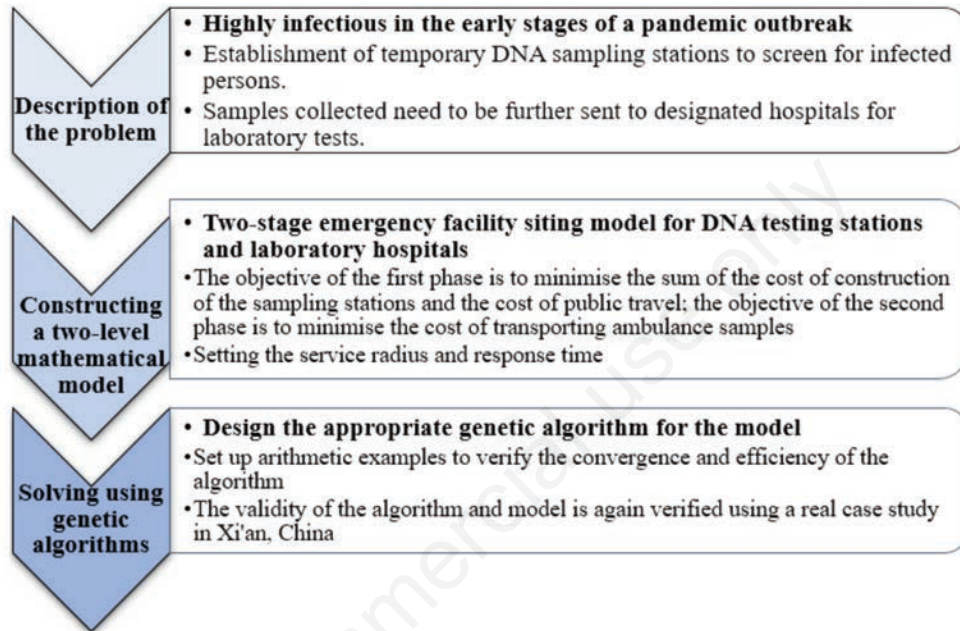
### Background

As a city with a population of 10 million, Xi'an faced a high risk of importing the COVID-19 infection. The establishment of a 15-minute DNA sampling circle, proposed and implemented by the Xi'an municipal government in 2022, would circumvent large-scale outbreaks and potential dangers on the one hand, and would also be beneficial in providing the public with more convenient healthcare services on the other. As the core area of Xi'an, Beilin District has a complex population flow, and in order to maximise the level of medical emergencies, Beilin District has established a relatively well-established COVID-19 testing service station in 2022. We therefore focused on the Beilin District of Xi'an as study area and defined 2022 as the study year.

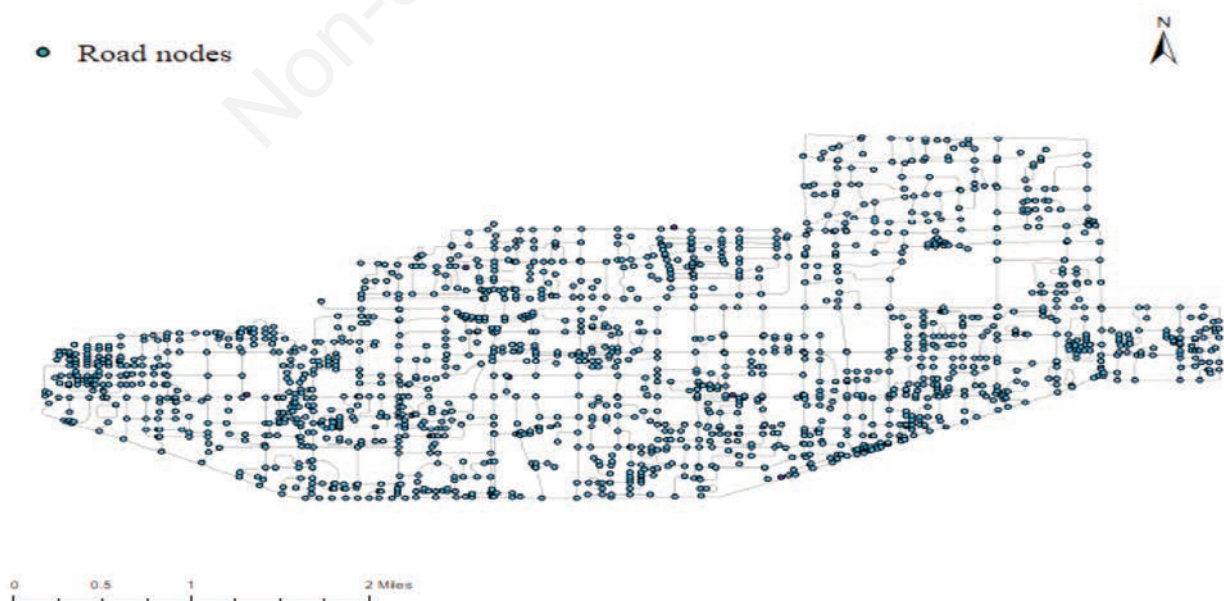
The road information of Beilin District was obtained from the Resources and Environmental Science and Data Center, Institute

of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences; population data and residence data for Xi'an were obtained from the 2021 Xi'an Statistical Yearbook, the Geoscience Data Network and the National Geographic Information Database (NGID). The intersections of secondary roads in Beilin District were used as demand points and candidate facility points (Figure 2). The demand points in the first stage are residential areas, representing the number of residents to be sam-

pled at that point, and the candidate facility points are DNA sampling stations, representing whether or not to construct a sampling station at the demand point (candidate facility point). Demand points and candidate facility points are the same set of points, *i.e.*, each demand point has the potential to be a sampling facility point, but due to resource constraints, different demand points (residential areas) may be assigned to the same facility point, and a sampling station can only be constructed at some of the demand points.



**Figure 1.** Flow chart describing what needs to be done during an epidemic.



**Figure 2.** Map of the study area with the stage I candidate sites.

Using ArcGIS software (ESRI, Redlands, CA, USA), the road network distances between points were firstly calculated to form distance matrix 1, while the number of residents was aggregated to the corresponding candidate points based on the proximity principle, as the demand for the candidate points in the first stage (see Table A in the Appendix).

The hospitals with medical testing qualifications announced by the Health Planning Commission of Beilin District, Xi'an (Table 1) were used as candidate sites for the second stage of the facility (Figure 3). The demand point of the second stage is the collection of sampling stations selected in the first stage, and the facility point is the location of the DNA laboratory, which are different sets of points, i.e., the samples collected at the sampling stations are transported to the laboratories that meet the constraints. The road network distances between the candidate sampling sites and the candidate DNA laboratories were calculated by using the ArcGIS software to form distance matrix 2.

### Model construction

#### General

A two-stage location model was created with aim of minimizing the number of sampling stations, DNA-laboratories and the transport of samples between them. We chose a complete coverage model including these facilities and the road network joining them based on existing facility requirements, integrated time constraints and transfer capacity. This first stage concerned the location of residential dwellings and the prevailing demand for testing with the walking time to giving the coverage radius. This was used to minimize the construction cost of sampling stations and the travel cost required for the residents to get there. The second stage consisted of minimizing the transport cost of samples from the sampling stations to the DNA-laboratories, where the sample preservation time was the constraint governing the service radius. Based on this a GA was designed.

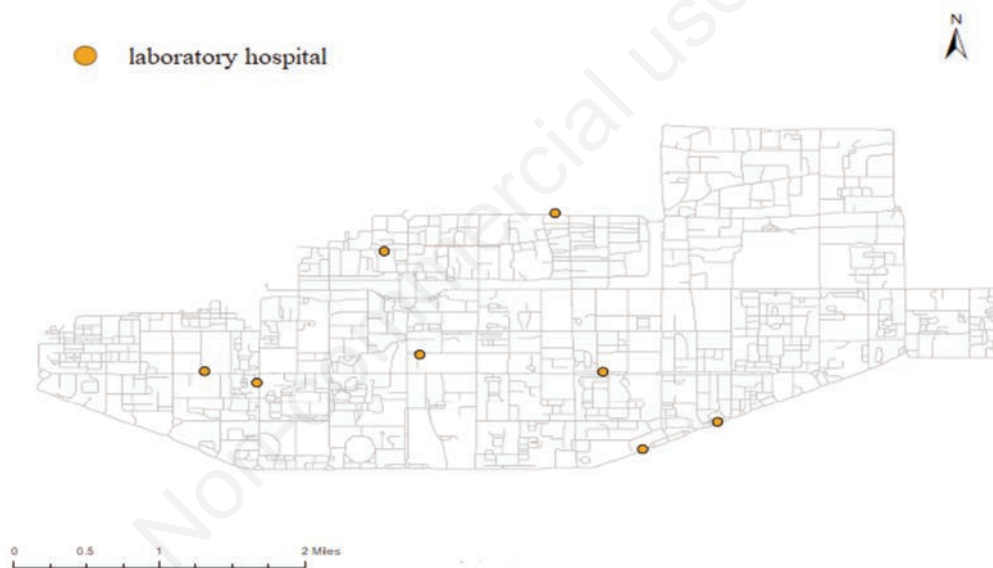


Figure 3. Map of the study area with the stage II candidate sites.

Table 1. Laboratories with DNA testing capability in Beilin District according to Xi'an Municipal Health Commission.

Hospital	Postal address
Xian NO.1 Hospital	30 Fen Lane, South Main Street
Xian NO.9 Hospital	151 East Section of South Second Ring Road
Honghui Hospital	555 Friendship East Road
Genertec Universal Medical Hospital	319 East Section of South Second Ring Road
Shaanxi Province Friendship Hospital	277 Friendship West Road
Shaanxi provincial people's Hospital	256 Friendship West Road
Beilin District Hospital of Traditional Chinese Medicine	38 East Sheep Market
Xijing 986 hospital	6 Construction West Road

### Sampling stations

The optimal location and number of sampling stations that would produce the minimum travel in addition to the cost of constructing new sampling stations were the objectives of the first stage. According to National Health Commission, the capacity of the sampling stations has a fixed value, here denoted as  $a_j$ , while the street network nodes were denoted as  $i = \{1, 2, \dots, I\}$ ; the number of residential areas around the nodes of the second-level street network were termed demand points and denoted as  $b_i$ ; the street network nodes in the secondary districts were taken as facility candidate points and denoted as  $j = \{1, 2, \dots, J\}$ ; unit construction costs as  $c_j$ ; the distances between the demand points and the facility points as  $d_{ij}$ ; the demand assigned to a candidate facility point as  $y_{ij}$ ; and the service radius of the sampling stations (set at 15-minute walking distance) as  $D_c$ .

### Assumptions

i) The limitations in demand and capacity at both demand points and candidate facility points are known; ii) The demand for testing does not exceed the capacity of the facility candidate point; iii) Demand can be covered by more than one facility point, i.e. residents can go to different sampling stations; iv) Distances are measured using the road network with residents reaching the sampling stations via primary and secondary roads, excluding lower-level roads, and only focusing on distance without considering effects of potential road works; and v) Residents walk at a uniform speed;

### DNA-laboratories

The facility candidate points for the location of DNA-laboratories in the second stage were denoted as  $k$ ; the distances between sampling stations and laboratories as  $s_{jk}$ ; shipment from the sampling stations to the facility point as  $u_{jk}$ ; according to official guidelines, medical reagents should be tested within a specified time to assure the effectiveness of the reagents used. This time limit should be the service radius of this stage and recorded as  $D_c'$ . Based on the efficiency of the healthcare facility, the number of people sampled per hour at each sampling point is  $n$ . The working time at the sam-

pling point is  $\frac{\sum_{i \in M_j} y_{ij}}{n}$ , from which the number of transportation

units can be calculated; in the second stage, the location and number of DNA-laboratories were chosen to minimize the transportation cost.

### Assumptions

i) The sampling stations have capacity limitations, while the DNA-laboratories do not, disregarding laboratory capacity and the time for reagents to reach laboratories; ii) Transportation costs are only related to the number of samples and distance, without considering other costs; and iii) The speed for reagent transport is constant, and the per-unit transportation cost is the same, without considering the impact of traffic congestion or road works.

### Variables

The parametric variables used in this study are given in Table 2.

### Sampling station construction costs

The network of sampling stations and the DNA-laboratories network must meet the stipulated security principles and need

therefore to be carefully distributed. The minimum value of the construction cost of a sampling station can be expressed as:

$$\text{Min} \sum_{j \in N_i} c_1 x_j \quad (\text{Eq. 1})$$

### Public travel

The service radius was set at the distance possible to cover in 15 minutes by foot and this travel cost without considering any other costs was expressed as:

$$\text{Min} \sum_{j \in J, k \in K} c_2 y_{ij} d_{ij} \quad (\text{Eq. 2})$$

### Sample transportation cost

The transportation cost of samples and reagents from the sampling stations to the DNA-laboratories and the number of times that test reagents need to be transported were calculated taking sample storage time and driving distance of the transport vehicle into account as follows:

$$\sum_{j \in N_i} y_{ij} = b_i, \quad \forall i \in I \quad (\text{Eq. 3})$$

where  $\left[ \frac{\sum_{i \in M_j} y_{ij}}{n g_k} \right]$  denotes the number of times the reagents needed

to be transported, The total transport cost was taken as the unit cost  $r$  multiplied by the distance travelled (with the distance doubled to account for the return of the transfer vehicle).

### Constraints

#### Feasibility domain constraint

The service radius was set at 15-minute walking time;  $N_i$  represents the collection of sampling stations that can cover the residential area  $i$  within the  $D_c$  range. When the demand point is  $i$ , the facility point is selected from the set that can cover  $i$  within 15 minutes and is less than or equal to the original facility point set.  $M_j$  represents a collection of residential areas that can cover the sampling stations  $j$  within the  $D_c$  range, and  $N'_j$  represents a collection of laboratories  $k$  that can cover the sampling stations  $j$  within the  $D_c'$  scope.  $M'_k$  represents a collection of sampling stations  $j$  that can cover DNA-laboratory  $k$  within the  $D_c'$  range.

#### Capacity limitation

The first requirement is to ensure that all demand points are covered. Since the capacity of the sampling stations is limited, the number of allocations cannot exceed the overall capacity.

$$\sum_{j \in N_i} y_{ij} = b_i, \quad \forall i \in I \quad (\text{Eq. 4})$$

$$\sum_{i \in M_j} y_{ij} \leq a_j x_j, \quad \forall j \in J \quad (\text{Eq. 5})$$



**Supply/demand balance**

The whole system is a supply and demand balancing system where all samples collected must be distributed to the DNA laboratories.

$$\sum_{k \in N'_j} u_{jk} = x_j, \quad \forall j \in J \tag{Eq. 6}$$

Based on the above analysis, the following model was constructed where equation 7 represents the first-stage objective of minimizing the cost of construction of sampling stations and the cost of travelling for residents; equation 8 represents the constraint that all DNA testing needs in residential areas must be covered; and equation 9 indicates that the number of samples collected by the sampling station  $j$  must not exceed its testing capacity and that sampling can only be carried out if this sampling station has been selected:

$$\text{Min } \sum_{j \in N_i} c_1 x_j + \sum_{j \in J, k \in K} c_2 y_{ij} d_{ij} \tag{Eq. 7}$$

$$\text{s. t. } \sum_{j \in N_i} y_{ij} = b_{ij}, \quad \forall i \in I \tag{Eq. 8}$$

$$\sum_{i \in M_j} y_{ij} \leq a_j x_j, \quad \forall j \in J \tag{Eq. 9}$$

Equation 10 represents the second-stage objective of minimizing the sample transport cost; while equation 11 represents the constraint that ensures that the selected sampling stations must all be covered (but only once); equation 12 represents the constraint that samples can be shipped from the sampling station to the DNA-laboratory, but only if the laboratory has been selected; and equation 13 indicates the range of values of the decision variables:

$$\text{Min } \sum_{j \in J, k \in K} 2 \left[ \frac{\sum_{i \in M_j} y_{ij}}{n G_K} \right] s_{jk} v_k r \tag{Eq. 10}$$

**Table 2.** Parametric factors and their definition.

Parameter	Definition
$i$	The residential area
$I$	The set of residential areas
$b_j$	The population of residential area $i$
$j$	The sampling station
$J$	the sampling station network
$a_j$	The capability limitation of sampling station $j$
$c_1$	The construction cost of a sampling station
$c_2$	Cost of travelling per unit distance for the public to the sampling station
$d_{ij}$	The distance between residential area $i$ and the sampling station $j$
$D_c$	The service radius
$M_0$	An infinite number
$N_i = \{j   d_{ij} \leq D_c\}$	The set of sampling stations that can cover residential area $i$ within the $D_c$ range
$M = \{j   d_{ij} \leq D_c\}$	The set of residential areas of the sampling stations $j$ can be covered within the range of $D_c$
$n$	The number of people that can be sampled at a sampling station per hour
$\frac{\sum_{i \in M_j} y_{ij}}{n}$	The time needed to collect all samples at sampling station $j$
$G_K$	The maximum time interval for transportation of collected samples to a DNA-laboratory
$\left\lceil \frac{\sum_{i \in M_j} y_{ij}}{n G_K} \right\rceil$	The minimum number of transports required by a sampling station in a limited time period rounded upward, if $>1$ , counted as 1
$r$	Transfer vehicle transport cost
$s_{jk}$	Distance between sampling station $j$ and laboratory $k$
$x_j$	{1, Select the sampling station $j$ 0, sampling station $j$ was not selected}
$y_{ij}$	The number of residents assigned to the sampling station $j$
$v_k$	{1, Select laboratory $k$ 0, Laboratory $k$ is not selected}
$u_{jk}$	{1, Sampling station $j$ selects laboratory $k$ 0, sampling station $j$ does not select laboratory $k$ }

$$\text{s. t. } \sum_{k \in N'_j} u_{jk} = x_j, \quad \forall j \in J \tag{Eq. 11}$$

$$\sum_{j \in M'_k} u_{jk} \leq M_0 v_k, \quad \forall k \in K \tag{Eq. 12}$$

$$x_j, u_{jk}, v_k \in (0,1), \quad y_{ij} \text{ integer} \tag{Eq. 13}$$

### Use of genetic algorithm

The genetic algorithm takes the coding of the decision variables as operator and performs a set of operations, such as encoding, cross-over, mutation and fitness (described below) in a probabilistic manner applying what is called an elite preservation strategy (also described below).

#### Chromosome encoding

The implementation of genetic algorithms requires a primary encoding as this affects the subsequent genetic operators directly and largely determines the efficiency of the evolution. Various encoding methods exist and as we were faced with multiple decision variables, actual number encoding was used since it helps to reduce computational complexity. Each variable corresponds to a specific gene location on the chromosome and the number of variables determines the code length. Two chromosomes are designed according to Eq. 13. In the first stage,  $x$  indicates whether a facility point has been selected and  $y$  the number of facilities allocated; thus when  $x$  takes the value of 1 for a given genotype, the corresponding  $y$  columns sum to  $>0$  (Figure 4), which indicates that the site should be selected for facility construction and that a demand assigned to it. The second stage runs likewise coding the decision variables  $u$  and  $v$  correspondingly.

The genetic algorithm starts the iteration with the initial population as the initial point and uses random numbers to generate  $N$  data strings, each of which represents an individual, who together form a population. In the coding session, the initialization of the  $Y$  chromosome needs to follow the  $X$  chromosome. The initialisation of the  $V$  chromosome also needs to follow the initialisation of the

$$\begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_{j-1} & x_j \\ y_{11} & y_{12} & \dots & \dots & y_{1j} \\ y_{21} & y_{22} & \dots & \dots & y_{2j} \\ y_{31} & y_{32} & \dots & \dots & y_{3j} \\ y_{41} & y_{42} & \dots & \dots & y_{4j} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ y_{(i-1)1} & y_{(i-1)2} & \dots & \dots & y_{(i-1)j} \\ y_{i1} & y_{i2} & \dots & \dots & y_{ij} \end{bmatrix}$$

Figure 4. Chromosome coding mode.

U chromosome.

#### Fitness

This is an important metric in heuristic algorithms representing the ability of an individual trait to adapt to its environment (and gain progress in the same way as in real life where successful organisms reproduce better). It is used to measure the possibility of an ‘individual’ to reach the optimal solution in an optimization computation. The fitness function for a model with constraints is constructed by adding the constraints as penalty terms to the objective function. Each time ‘individuals’ are chosen from a population, the better-adapted ones are selected to enter the ‘offspring’ population. This operation is repeated until the population size becomes the same as the original one. We used a binary tournament selection method, which involves taking two ‘individuals’ from the population at once and selecting the better one to be retained in the next-generation based on comparison of fitness functions.

#### Cross-over

A two-point, cross-over approach was used, which refers to randomly setting two cross-over points in individual chromosomes and exchanging some ‘genes’ there. The procedure is as follows: two individuals were randomly selected from the parental generation (1, N) (where N is the length of the gene population on the chromosome) to generate two random numbers. These two random numbers determine the crossover position, where the two chromosomes are cut and then re-combined cross-wise generating two new ‘individuals’ as shown in Figure 5.

#### Mutation

This is done to determine whether ‘individuals’ should be mutated to maintain the diversity of the population needed to prevent the population from falling into a local optimum. This study used the uniform variation method, which first determines the location of an individual gene variation, then replaces the original

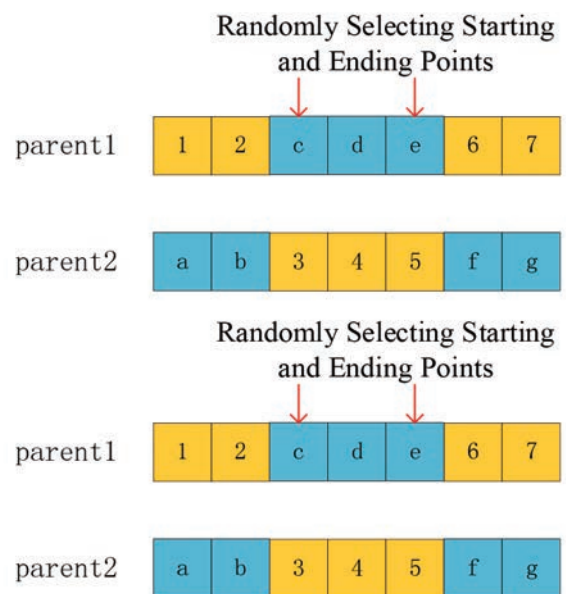


Figure 5. Two-point, cross-over schematic diagram.

'gene' value with a random number that conforms to a uniform distribution within a specific range according to a certain probability. For example, if  $X_k$  is the point of mutation in individual  $X = \{X_1, X_2, \dots, X_L\}$  that takes the value range  $[U_{min}, U_{max}]$  after a uniform mutation of the individual, the new individual  $X = \{X_1, X_2, \dots, X'_k, \dots, X_L\}$ , where  $X'_k = U_{min} + r(U_{max} - U_{min})$  appears with  $r$  being a random number that conforms to a uniform distribution over a range.

**Elite preservation strategy**

After the offspring with cross-variation had been obtained and merged with the parents to form a new population generation with a size of  $2N$ , the fitness is calculated and sorted, with the first  $N$  'individuals' with better fitness retained as the new population size. This elite retention strategy was adopted to optimize the population further. The flow of the designed algorithm is as follows: i) Initialize a population consisting of  $N$  individuals according to the coding rules; ii) Calculate the fitness of the current population; iii) Stop when the stopping condition is met; otherwise continue; iv) Independently select  $N$  individuals from the current population; v) Independently perform a cross-over operation in these individuals; vi) Independently mutate the  $N$  individuals after the cross-over operation; vii) Merge the parent and population resulting from the crossover mutation to obtain a population of size  $2N$ ; viii) From the merged population, select  $N$  individuals according to the selection algorithm to obtain a new generation population; and then ix) Return to step 2. The flowchart is shown in Figure 6.

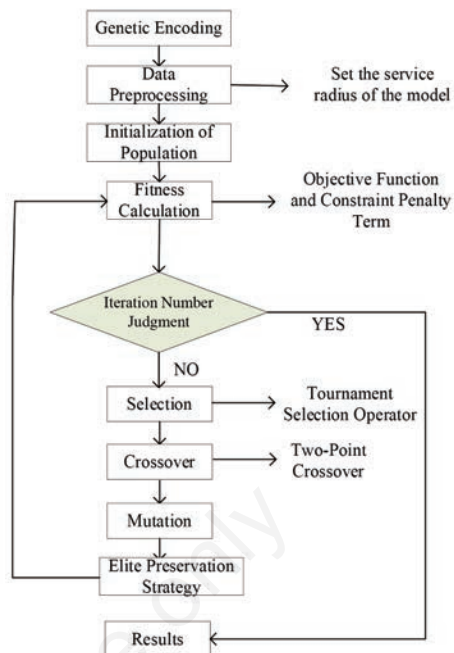


Figure 6. Flowchart of genetic algorithm design.

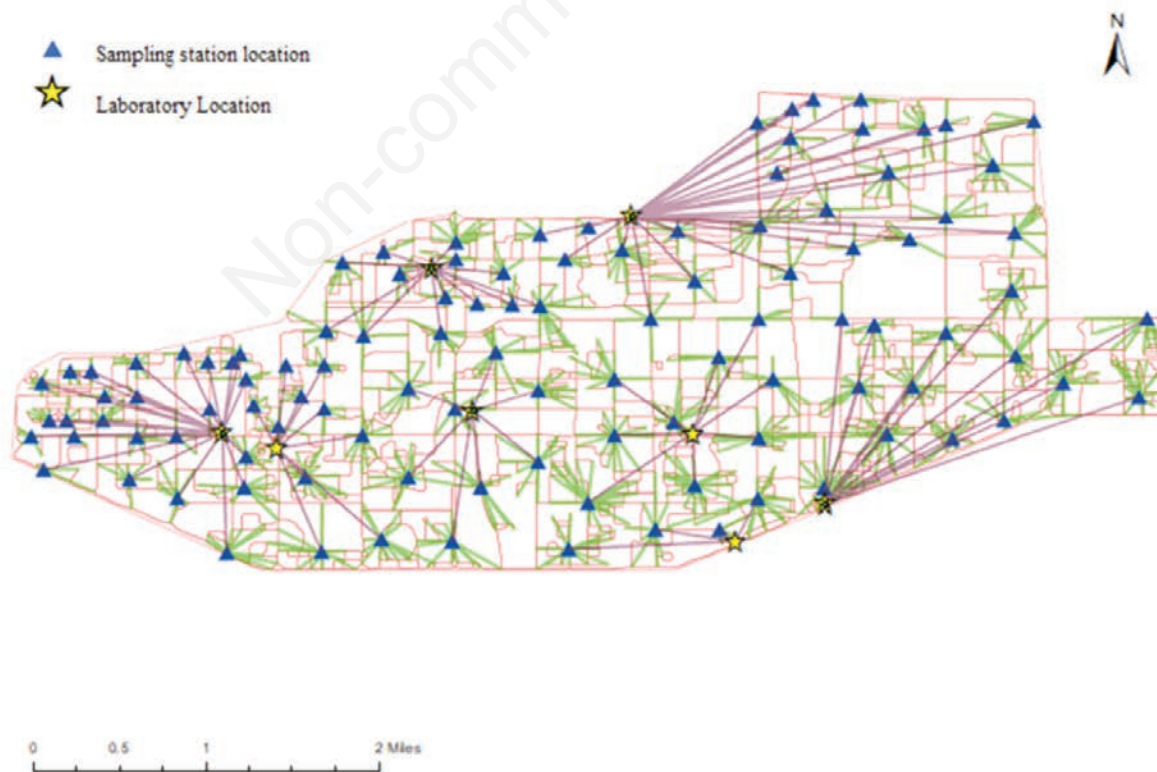


Figure 7. Location of sampling stations and DNA-laboratory in Beilin District.



### Algorithm testing

To evaluate the effect of a genetic algorithm, the algorithm should be tested with different indicators, mainly including the convergence index of the solution time and the solution. We used Gurobi (Kratica *et al.*, 2014; Saghand *et al.*, 2022) to generate five computational examples randomly with the cross-over probability ( $P_c$ )=0.7, the mutation probability ( $P_m$ )=0.05, the population size (POP)=100 and the iteration number  $G=50$ .

### Study variable settings

Each healthcare worker can collect about 360 samples per hour. Calculating on the basis of 5 healthcare workers for each sampling station, 1,800 samples can be taken at each sampling station per hour, and according to the “Guidelines for the Implementation of Regional Medical Testing for Novel Coronaviruses” of the National Health and Wellness Commission,

the sampling should be completed within 6 hours, and each sampling station can accommodate the collection of samples from 10,800 people. Assuming that each user walks 2,000 meters in 15 minutes, the first stage’s service radius was set at this length. Each sampling station costs 20,000 Chinese yuan (RMB) to build and 1 RMB for a residential trip. According to the regulations, follow-up samples must be tested within 4 hours and collected and transported every 30 minutes to 2 hours so the service radius of the second stage was set at 2 hours; the cost of the transfer vehicle within 3 km is 10 RMB and the unit transportation cost  $r$  was set at 10 RMB. Specific parameter values are shown in Table 5.

### Algorithm solving

The distance matrix, target model, constraints and parameters were entered into the genetic algorithm, and the language of the genetic was written by a Python application, which selected a total of 109 sampling stations and 8 laboratories.

**Table 3.** Comparison between the genetic algorithm and Gurobi.

Population size (pcs)	Algorithm solution time (sec)	Algorithm solution results	Gurobi solution time (sec)	Gurobi solution results
10*10	2.20	552936	3.99	551894
20*20	4.80	863242	34.50	861924
30*30	7.28	1183355	186.29	1180703
50*50	12.21	18578993	1918.19	18547776
100*100	35.53	35748325	7100.28	35626523

**Table 4.** Efficiency of the Genetic algorithm at different scales.

Population size(pcs)	Iterations (no.)	Solution time (sec)
10*10	50	2.75
10*10	100	6.21
100*100	50	36.72
100*100	100	80.63
200*200	100	229.77

**Table 5.** Relevant parameters.

Parameter	$D_c$	$n$	$r$	$G_K$	$C_1$	$C_2$
Numerical setting	2,000	1,800	10	2	20,000	1

$D_c$ , service radius;  $n$ , number of people sampled per hour at the sampling station;  $r$ , transfer vehicle transport cost;  $G_K$ , maximum time interval for transportation of collected samples to a DNA-laboratory;  $C_1$ , construction cost of a sampling station;  $C_2$ , cost per unit distance.

**Table 6.** Computational results of the locations of sampling ststion and DNA-laboratories.

Sampling station (serial number)	Laboratory
724, 27, 36, 55, 66, 67, 90, 99, 101, 122, 130, 132, 135, 161, 164, 165,180, 185, 198, 207, 215, 218, 222, 224	Shaanxi Province Friendship hospital
236, 287, 302, 325, 333, 353, 361,364, 416, 458	Shaanxi provincial people’s hospital
369, 385, 429, 459, 486, 554, 563,598, 599, 626, 682, 689, 724	Xian No.1 Hospital
497, 500, 571, 586, 635, 673, 713,716,	Honghui Hospital
727, 767, 812, 869, 915, 967, 997, 1067, 1086, 1118, 1140, 1143, 1149,	Beilin District Hospital of Traditional
1179, 1206, 1243, 1265, 1272, 1332, 1367, 1394, 1422, 1426, 1476, 1503, 1536	Chinese Medicine
808, 851, 856, 957, 991, 1031, 1077, 1080, 1112	Xijing 986 hospital
1199, 1229, 1259, 1295, 1323, 1373, 1418, 1432, 1483, 1499, 1505, 1565, 1607, 1620	Xian NO.9 hospital
771, 930, 1032, 1069	Genertec Universal Medical

## Results

### Algorithm testing

Table 3 shows the difference between the genetic algorithm and Gurobi with respect to the time needed to solve the equations. It can be seen that the genetic algorithm is faster than Gurobi and that it varies with the scale. In addition, the efficiency of the genetic algorithm used in this study was also progressively more stable with increasing scale. The optimal solution obtained by the genetic algorithm was close to that obtained by Gurobi. To further prove the algorithm's effectiveness, the number of demand points and facility points as well as the number of iterations were tested with unfading results. Indeed, it was found that the solution time of the model grows as the scale of the independent variables increases and the effect of the scale on the solution efficiency is greater than the number of iterations (Table 4). Thus, the results converge faster at the different demand points and facility candidate point sizes, which is evidence that this algorithm is sufficiently good.

### Comparative analyses

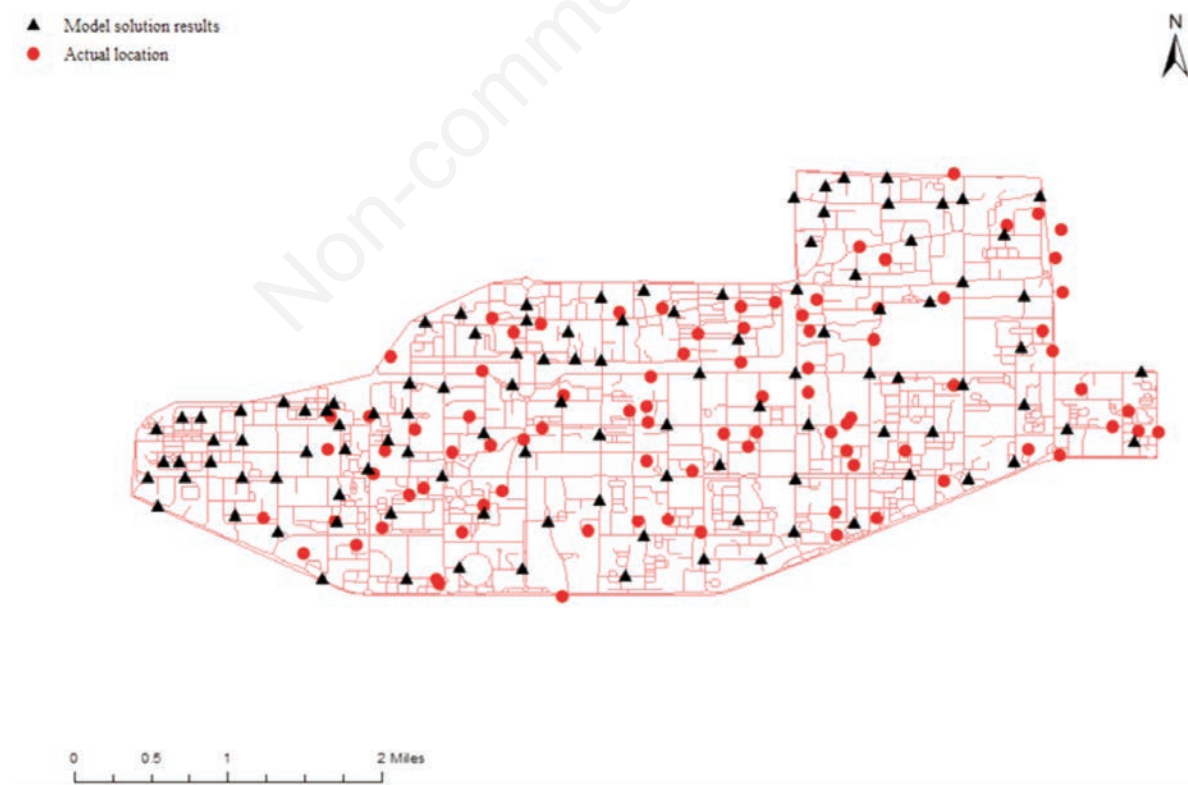
The results obtained by the algorithm were compared with the location of sampling stations and DNA-laboratories in Beilin District during the COVID-19 epidemic and visualized with ArcGIS software. As can be seen from Figure 8, the results of the model constructed in this study are roughly evenly distributed in each region, covering the entire area and many sampling stations

are close to the actual positions (Figure 8), all laboratories with DNA testing capabilities were selected (Figure 7), which indicates that the model and algorithm are effective. At the same time, considering the construction cost, the location scheme proposed can cover all existing needs and satisfy the sampling time constraints, often providing a more comprehensive location decision. It is worth noting that the results obtained by the model selected 20 more sampling stations than the actual situation, and with regard to sampling stations, there were often few, sometimes none located in some areas in practice. This is because the model is based on the location of facilities under the condition of sufficient need for DNA-testing, but as unable to organize staff to handle the situation when the pandemic first struck. Only selective construction could be carried out in some areas.

## Discussion

### General

This paper presents the construction of a two-stage covering location model considering capacity constraints and rapid response. While ensuring that all demands are met, the model aims to minimize costs associated with transportation, building of facilities and residents' travel. Taking into account the capacity constraints of the sampling stations and the given service radii, a two-stage project covering residents and sampling stations on the one hand and the transport to and location of laboratories with capacity



**Figure 8** Comparison between computational results of sampling stations and actual locations in Beilin District 2022.

to perform DNA testing on the other. This provides a novel approach for emergency facility location, more aligned with real-world scenarios. In terms of model solution, we designed a genetic algorithm whose solution results were compared with those obtained using the Gurobi algorithm package. The genetic algorithm designed using the Python language outperformed the Gurobi algorithm in terms of speed even when dealing with large the population sizes and time constraints. Taking the Beilin District of Xi'an City as example, 1,636 road nodes were selected as the first-stage demand points and candidate testing points, with 8 medical institutions meeting national medical testing qualifications chosen as the second-stage candidate laboratory points. The results of the algorithm solution show that 109 sampling stations need to be constructed in the first stage, and all the laboratory candidate sites for the second stage have been selected.

The results of the model constructed in this study show that proposed localisation of facilities are roughly evenly distributed in each region, covering the entire area, with many alternative points close to the actual positions (Figure 8), often providing a more comprehensive location decision than the actual one. Considering the costs of building extra sampling stations and DNA-laboratories, citizens' travel and sample transportation, the location scheme proposed here meets the capacity of the facilities and the time requirements. It is worth noting that the results obtained by the model selected 20 more sampling stations than the current scheme and there were often few, sometimes no DNA-laboratories located in some areas in practice. This is because we designed the model based on the location of facilities under the condition of sufficient need for DNA-testing, but did not include organization of staff to handle the situation when the pandemic first struck.

Existing research has demonstrated the critical importance of sampling stations (Zhao *et al.*, 2021) and nucleic acid screening stations (Liu *et al.*, 2022 Niu *et al.*, 2024) in preventing the spread of health crises such as the COVID-19 pandemic and in reducing their profiles. Recent studies have also begun to address issues regarding the allocation of healthcare workers (Saidani *et al.*, 2021) and subjects (Xie *et al.*, 2022) at sampling stations, but little research has been done on the siting of sampling stations and DNA laboratories, and this paper aims to fill this gap.

Firstly, the research model can provide specific decision-making suggestions for the national emergency management department. Models similar to the one proposed here can be constructed by setting different parameters according to the situation at hand, such as locations for candidate facilities, with reasonable planning arrangements provided for resource allocation in emergency response or national disaster prevention. At the same time, it helps to improve the collaborative response ability between the relevant departments and provides academic support for the formulation of rules and regulations of emergency management systems.

Secondly, the research model can provide feasible solutions for the location of medical testing units in relation to hospitals and other health care units. These units set the coverage radius and limit capacity of the facility points according to the location characteristics. The application of models as the one described here can minimize the cost of facility locations, improve the utilization rate of scarce resources and reduce construction costs. The model established in this study is thus universal and can be applied in different fields where time requirements play a role, *e.g.*, centres for logistics and distribution.

Thirdly, the theoretical model can provide residents with corresponding plans to deal with future emergencies. Combining with

relevant requirements and residents' reality, setting different coverage service radius, and adopting the location mode with the goal of minimizing the construction cost of the test site and the travel cost of residents, can guide cooperation between residents and government enterprises.

### Limitations of the study

Due to the limited level, this study is still deficient in many places and future research can consider improving it from the following perspectives: i) Only the optimization problem of cost and number of facilities was considered, input parameters were all set, and a single time node selected to analyze the problem, without considering multiple choices of citizens and dynamic changes of demand points; ii) With regard to the cost factor, this paper only considered the average construction cost of the DNA-laboratories and not the impact of various other factors, which need to be further explored in the subsequent research; iii) The genetic algorithm was designed to solve the problem, but the efficiency of solving large-scale problems needs was not compared with other heuristic algorithms; iv) The research is based on the relevant data provided by the National Health Commission of China. However, in the actual situation, the location options and capacity limitations of medical testing enterprise outlets can be different from these data. The research on the location of sampling stations only remains at the theoretical level, and many unforeseen factors can be expected in actual operative situations.

### Conclusions

The research bears significant implications for the location of emergency facilities following sudden incidents and can provide effective decision-making guidance for the location of medical testing facilities. The results indicate that 109 medical sampling stations and 8 DNA-laboratories need to be established. Comparing with the actual situation, the results validate the effectiveness of the model and algorithm constructed and discussed in this paper.

### References

- Ahmadi-Javid A, Seyedi P, Syam S, 2016. A survey of healthcare facility location. *Comput Oper Res* 79:223-63.
- Alizadeh RNT, Nishi T, 2020. Hybrid set covering and dynamic modular covering location problem: Application to an emergency humanitarian logistics problem. *Appl Sci* 10:7110.
- Amideo AE, Scaparra M P, Sforza A, 2021. An integrated user: system approach for shelter location and evacuation routing. *Networks* 78:46-68.
- Bashiri M, Rezanezhad M, Tavakkoli-Moghaddam R, Hasanzadeh H, 2018. Mathematical modeling for a p-mobile hub location problem in a dynamic environment by a genetic algorithm. *Appl Math Model* 54:151-69.
- Bélanger V, Lanzarone E, Nicoletta V, 2020. A recursive simulation-optimization framework for the ambulance location and dispatching problem. *Eur J Oper Res* 286:713-25.
- Caunhyea AM, Zhang Y, Li M, Nie X-F, 2016. A location-routing model for prepositioning and distributing emergency supplies. *Transp Res E Logist Transp Rev* 90:161-76.



- Davis HE, McCorkell L, Vogel JM, Topol EJ, 2023. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol* 21:133-46.
- Dong S, Deng Y, 2018. Optimization model of emergency station location for urban rail transit. *IOP Conf Ser Earth Environ Sci* 189:062006.
- Farrokhi-Asl H, Tavakkoli-Moghaddam R, Asgarian B, Sangari E, 2017. Metaheuristics for a bi-objective location-routing-problem in waste collection management. *J Ind Prod Eng* 34:239-52.
- Gendreau M, Laporte G, Semet F, 1997. Solving an ambulance location model by tabu search. *Location Sci* 5:75-88.
- Gu Zhou Y, Das A, 2018. Medical relief shelter location problem with patient severity under a limited relief budget. *Comput Ind Eng* 125:720-8.
- Karimi H, Bashiri M, 2011. Hub covering location problems with different coverage types. *Scientia Iranica* 18:1571-8.
- Kraticek J, Dugošija D, Savić A, 2014. A new mixed integer linear programming model for the multi level uncapacitated facility location problem. *App Math Model* 38:2118-29.
- Liu CJ, Yang CY, Chen YL, Wang SS, Chu CM, Hsieh MH, Huang SJ, Chang YP, Hsu CC, 2022. Prompt successful response to a COVID-19 outbreak: Performance of community-based rapid screening station. *J Formosa Med Assoc* 121:2356-2359.
- Liu J, Li Y, Li Y, Chen Z., Lian X., Zhang Y, 2023. Location optimization of emergency medical facilities for public health emergencies in megacities based on genetic algorithm. *Eng Construct Arch Manag* 30:3330-56.
- Liu Y, Li Z, Liu J, 2016. A double standard model for allocating limited emergency medical service vehicle resources ensuring service reliability. *Transp Res C: Emerg Technol* 69:120-133.
- Liu Z, Li Z, Chen W, Zhao Y, Yue H, Wu Z, 2020. Path optimization of medical waste transport routes in the emergent public health event of covid-19: a hybrid optimization algorithm based on the immune-ant colony algorithm. *Int J Env Res Pub He* 17:1-18.
- Long S, Zhang D, Li S, Li S, 2023. Two-stage multi-objective stochastic model on patient transfer and relief distribution in lockdown area of COVID-19. *Int J Environ Res Public Health* 20:1765.
- Mantzaras G, Voudrias EA, 2017. An optimization model for collection, haul, transfer, treatment and disposal of infectious medical waste: application to a Greek region. *Waste Manag* 69:518-34.
- Murali P, Ordóñez F, Dessouky MM, 2016. Modeling strategies for effectively routing freight trains through complex networks. *Transp Res C: Emerg Technol* 70:197-213.
- Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, Agha M, Agha R, 2020. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *Int J Surg* 78:185-93.
- Nikzami M, Baradaran V, 2020. A healthcare logistic network considering stochastic emission of contamination: Bi-objective model and solution algorithm. *Transp Res E: Logist Transp Rev* 142:102060.
- Niu Y, Cheng Y, Chan C, Chaou C, Yen C, Fang C, 2024. SARS-CoV-2 rapid antigen testing positive rate in community testing stations as an indicator for COVID-19 epidemic trend, Taipei, Taiwan, May to August 2021. *J Formosa Med Assoc* 123:716-9.
- Nogueira L, Pinto L, Silva P, 2016. Reducing emergency medical service response time via the reallocation of ambulance bases. *Health Care Manag Sci* 19:31-42.
- Oksuz, MK, Satoglu SI, 2020. A two-stage stochastic model for location planning of temporary medical centers for disaster response. *Int J Disaster Risk Reduction* 44:101426.
- Pichka K, Bajgirani AH, Petering M, 2018. The two echelon open location routing problem: Mathematical model and hybrid heuristic. *Comput Ind Eng* 121:97-112.
- Rabbani M, Heidari R, Farrokhi-Asl H, Rahimi N, 2018. Using metaheuristic algorithms to solve a multi-objective industrial hazardous waste location-routing problem considering incompatible waste types. *J Clean Prod* 170:227-41.
- Saghand P G, Charkhgard H, 2022. Exact solution approaches for integer linear generalized maximum multiplicative programs through the lens of multi-objective optimization. *Comput Oper Res* 137:105549.
- Saidani M, Kim H, Kim J, 2021. Designing optimal COVID-19 testing stations locally: A discrete event simulation model applied on a university campus. *PLoS One* 16:e0253869.
- Sanci E, Daskin MS, 2019. Integrating location and network restoration decisions in relief networks under uncertainty. *Eur J Oper Res* 279:335-50.
- Schneeberger K, Doerner KF, Kurz A, 2016. Ambulance location and relocation models in a crisis. *Cent Eur J Oper Res* 24:1-27.
- Seraji H, Tavakkoli-Moghaddam R, Asian S, 2021. An integrative location-allocation model for humanitarian logistics with distributive injustice and dissatisfaction under uncertainty. *Ann Oper Res* 319:211-57.
- Siddiq AA, Brooks S C, Chan T C Y, 2013. Modeling the impact of public access defibrillator range on public location cardiac arrest coverage. *Resuscitation* 84:904-9.
- Wang H, Ma X, 2021. Research on multiobjective location of urban emergency logistics under major emergencies. *Math Probl Eng* 5577797:12.
- Wang Y, Zhu X, 2022. A multi-regional collaborative optimization model of emergency medical materials for responding to COVID-19. *Processes* 10:1488.
- Xie N, Liu Z, Chen X, Li S, 2022. Fair assignment for reserved nucleic acid testing. *Sustainability* 14:11752.
- Zhao H, He Y, Brister F, Yang L, Li G, Ling Y, Ying Y, 2021. How Can Nursing Teams Respond to Large-Scale COVID-19 Screening? *Front Public Health* 9:681255.
- Zhong S, Cheng R, Jiang Y, 2020. Risk-averse optimization of disaster relief facility location and vehicle routing under stochastic demand. *Transp Res E: Logist Transp Rev* 141:102015.
- Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Chen H, Cao B, 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395:1054-62.