Context-aware Heatstroke Relief Station Placement and Route Optimization for Large Outdoor Events

Yan Wu
Xi'an Jiaotong University
Tianqi Xia (xiatianqi@csis.u-tokyo.ac.jp)
The University of Tokyo: Tokyo Daigaku https://orcid.org/0000-0003-4575-6883
Adam Jatowt
National Institute of Advanced Industrial Science and Technology
Haoran Zhang
The University of Tokyo - Kashiwa Campus: Tokyo Daigaku - Kashiwa Campus
Xiao Feng
Xi'an Jiaotong University
Ryosuke Shibasaki
The University of Tokyo - Kashiwa Campus: Tokyo Daigaku - Kashiwa Campus
Kyoung-Sook Kim
National Institute of Advanced Industrial Science and Technology

Research

Keywords: Optimization, context-aware, pedestrian flow, Olympic Games

Posted Date: February 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-219494/v1

Title Page

Context-aware Heatstroke Relief Station Placement and Route Optimization for Large Outdoor Events

Yan Wu[1,2], Tianqi Xia[*1,3], Adam Jatowt[3,4], Haoran Zhang[1], Xiao Feng[2], Ryosuke Shibasaki[1], Kyoung-Sook Kim[3]

[1] The University of Tokyo. 277-0882 Kashiwanoha 5-1-5, Kashiwa, Japan

[2] Xi'an Jiaotong University. 710049 28 Xianning West Road, Xi'an, Shaanxi, China

[3] National Institute of Advanced Industrial Science and Technology. 135-0064, Aomi, Koto, Tokyo, Japan

[4] Kyoto University. 606-8501, Yoshida-honmachi, Sakyo-ku, Kyoto, Japan

Corresponding author: Tianqi Xia;

Email: xiatianqi@csis.u-tokyo.ac.jp

Context-aware Heatstroke Relief Station Placement and Route Optimization for Large Outdoor Events

Yan Wu[1,2], Tianqi Xia[*1,3], Adam Jatowt[3,4], Haoran Zhang[1], Xiao Feng[2], Ryosuke Shibasaki[1], Kyoung-Sook Kim[3]

[1] The University of Tokyo. 277-0882 Kashiwanoha 5-1-5, Kashiwa, Japan

[2] Xi'an Jiaotong University. 710049 28 Xianning West Road, Xi'an, Shaanxi, China

[3] National Institute of Advanced Industrial Science and Technology. 135-0064, Aomi, Koto, Tokyo, Japan

[4] Kyoto University. 606-8501, Yoshida-honmachi, Sakyo-ku, Kyoto, Japan

Abstract: Heatstroke is becoming an increasingly serious threat to outdoor activities, especially, at the time of large events organized during summer, including Olympic Games or various types of happenings in amusement parks like Disneyland or other popular venues. The risk of heatstroke is naturally affected by a high temperature, but it is also dependent on various other contextual factors such as the presence of shaded areas along traveling routes or the distribution of relief stations. This study proposes a framework to evaluate the heatstroke risk in geographical areas by utilizing context-aware indicators which are determined by large and heterogeneous data including facilities, road networks and street view images. Based on the evaluation metric of the total heatstroke risk in the target area, we propose a Mixed Integer Nonlinear Programming model for optimizing routes of pedestrians, determining the location of relief stations and the supply volume in each relief station. Our experiments conducted on the planned site of the Tokyo Olympics and simulated during the two weeks of the Olympics schedule indicate that planning routes and setting relief stations with our proposed optimization model could effectively reduce heatstroke risk, thus letting organizers better prepare for the event.

Keywords: Optimization, context-aware, pedestrian flow, Olympic Games

1. Introduction

With global warming and heat island effects caused by urbanization, heatstroke is becoming an increasingly severe threat to outdoor activities in summer. On the other hand, summer is a popular season for outdoor trips due to summer vacations and many events that normally take place in this

season. In such a case, it is fundamental to reduce heatstroke risk when holding large outside events, especially for pedestrians who are likely to be exposed to high temperatures.

There are several strategies that local governments or event holders could take to reduce heatstroke during large outdoor events. The easiest one is to set outdoor events to an earlier or later time of a day. However, this strategy would potentially reduce the number of visitors for early events, or increase the cost and other risks for night time events. Considering the potential income loss and cost increase, it is usually more reasonable for event organizers to minimize the risk in other ways such as providing relief stations to give first-aids to those who suffer or may suffer from heatstroke, as well as carefully planning routes that pose the smallest risk. However, simply implementing either of these two strategies could be rather challenging when the walkable space is complicated, e.g., spanning a large area with multiple origins and destinations (OD) and with complex routes between many OD pairs. In such a case, setting relief stations for all the routes would translate to a high cost. On the other hand, though heatstroke risk could be reduced by enforcing visitors to choose the optimal routes predefined by the organizers, heatstroke risk on long routes could be still inevitably high. To prevent these scenarios, setting up temporal relief stations based on predefined routes would ensure safer travel and at the same time keep the cost under control.

On the other hand, the location optimization scenario of heatstroke prevention varies from other similar scenarios in that heatstroke risk is sensitive to different environmental contexts within the walkable space. Specifically, a road segment with a better physical environment (e.g., an environment with less solar radiation) might pose less heatstroke risk than the road segment with a poor physical environment. As a result, pedestrians walking along the routes with high heatstroke risk cause high potential demand for assistance such as providing shelter or water. However, the traditional data used for representing spatial contexts, such as satellite images and statistic information, are usually characterized by low spatial resolution that cannot be applied to represent the detailed environmental context. As a result, the existing studies usually have the Uncertain Geographic Context Problem (UGCoP)[1] of ignoring health risk variance in different spatial and temporal units.

The development of big data and location-based services (LBS) makes however detailed contextual information available for reasoning about the health risk of diverse locations. Different

from the traditional data collection and processing approaches, contextual information can be now collected through road network-based sensors such as cars and street cameras. The data provided thanks to the deployment of such sensors can help to distinguish contextual differences in road network when it comes to health risks such as heatstroke. Based on large contextual data, our study proposes a problem of optimizing routes and placement of heatstroke relief stations in a road network within walkable spaces. The objective of this problem is to minimize global heatstroke risk, which is calculated by microscope contextual information on each road segment. Specifically, we propose a heatstroke risk model that measures the heatstroke risk of each road segment with different indicators calculated using heterogeneous data including characteristics of road segments. Based on this model, we conduct a case study in a real-world scenario of Tokyo Olympic Games with heterogeneous data collected from different data sources including Olympic schedules, facility locations and road network information. With the calculated contextual heatstroke risk, we further propose a Mixed Integer Nonlinear Programming (MINLP) problem to optimize pedestrian routes, relief station locations and supply volume of each station at different time.

The contributions of this study can be listed as follows:

(1). We propose a novel emergency service problem that can be applied in large outdoor event scenarios with multiple walking flows.

(2). We introduce a dedicated framework to solve this task. Both supply and demand are considered during the facility optimization. Specifically, the demand in this study is represented by pedestrian flows instead of demand points.

(3). The model we propose does not only plan the placement of relief stations but also determines the optimal routes for the dynamic pedestrian flows.

2. Related Work

2.1 Emergency facility location optimization

Emergency facilities are of great significance in public health as they can provide first aid to emergency victims to reduce casualties. Comparing to non-emergency facilities, the demand for emergency facilities is more time-sensitive and dependent on particular emergency scenarios. Therefore, although the existing studies on emergency facility location optimization (EFLO) have

similar objectives with other FLO problems to maximize covering location (MCLP)[2] or minimize cost or the number of facilities while making sure that the entire target region remains covered[3], the existing studies distinguish themselves from other EFLO problems in terms of optimization targets and problem settings[4].

From the perspective of optimization targets, several types of emergency facilities have received attention in existing studies. Off-site public access devices (OPAD) usually refer to those facilities that provide medical service out of regular healthcare facilities, e.g., automated external defibrillator (AED). Siddiq et al.[5] pointed out that the limited accessibility, poor visibility and lack of registration could influence AED demand and set different coverage values for different devices in location optimization.

Emergency center or department is another type of common emergency facility. Different from the OPAD devices, each emergency center usually has a larger supply volume and higher cost. Thus, in addition to studies of setting permanent emergency centers or departments with coverage problem settings[6], a lot of studies have been conducted on optimizing locations for temporary relief emergency centers. In these studies, emergency medical service demand, supply and accessibility distribution is supposed to vary in different emergency scenarios and their phases[7]. Schempp et al.[8] proposed a framework of utilizing SNS data to detect the emergency demand distribution and optimize the temporal rescue centers via global particle swarm optimization and mixed-Integer linear programming. Oran et al.[9] proposed a location-routing problem that considers the propriety of the locations and solves the problem using an mix integer programming solver.

Finally, ambulance transportation is significant in EMS and has received much attention in some existing studies. Comparing to other emergency facilities, the optimization targets of ambulance transportation are not limited to the location of ambulance stations, but also include the relocation and dispatching of ambulances. Since the supply volume of an ambulance fleet is limited, a problem that can represent the vacancy of ambulances is necessary for real-world application. This problem could be either solved by deterministic models via the backup of multiple ambulances to cover EMS demands[6], or solved by the probabilistic models that represent the access information by the probability of ambulance vacancy[10]. Daskin[11] proposed a maximum expected coverage location problem (MEXCLP) that the expected coverage of ambulances is calculated by the vacant

probability. The MEXCLP could be improved via short-term dynamic settings of ambulance supply and demand[12].

The problem approached in this study distinguishes itself from the existing EFLO problems in the following aspects: first, we assume the potential "patients" of our problem are pedestrians; second, the relief stations are not the destinations of the pedestrians, and the heatstroke risk is generated during the trip. Altogether these differences make our task a novel research problem in the field of emergency service.

2.2 Context-aware LBS application

Context-aware LBS application refers to those LBS applications that can provide service based on their present context including location, time and companions[13]. This extra information is of great significance to application users as their contexts vary from time to time, and any analysis with uncertain contextual information will generate bias and reduce application utility[1].

The last decade has witnessed the great development of contextual-aware LBS applications due to the availability of spatial and temporal data in high resolution[14]. A common context-aware implementation in the LBS application is to recommend points of interest (POI) to the visitors based on their spatio-temporal information, profiles and historical records. Yao et al.[15] proposed a tensor-factorization-based recommender system to recommend POIs with multi-dimensional contextual information. Besides single POI, several studies focused on recommending POI sequences. Chen and Jiang[16] proposed a context-aware personalized POI sequence recommendation system to recommend a sequence of POIs via reinforcement learning. Laß et al.[17] represented POIs by a graph and incorporated contextual information including historical records and traveling time into the traditional two-dimension user-item recommender system.

Another important implementation of context-aware information in LBS applications is route recommendation and navigation. In routing applications, contextual information could be utilized to measure the quality of each road segment to improve the traditional routing application by providing users with scenic, safe or attractive routes among other dimensions[18]. Specifically, the contextual information is utilized to evaluate each road segment based on its attractiveness or risk and choose the roads which are more attractive or less dangerous for different application scenarios. Attractiveness is usually evaluated by the accessibility to POIs[19] or the landscape diversity[20]

while risk can be assessed by social environment represented by metrics such as accidents, crimes and population density[21, 22], or physical environment such as solar radiation and infrastructure preparedness[23]. Generally, the contextual information could be collected from official statistical data[24], location-based social networking (LBSN) platforms[21], or web map services[25]. Unlike the above-mentioned researches focused on recommending individual optimal routes, this study focuses on recommending routes for a group of people with the global optimal objective.

3. Method

3.1 Problem definition and setting

Let us assume there is a planned large, long-lasting event that consists of several sub-events to be held in a given area. The area is composed of a road network with multiple venues, hotels, train stations and scenic spots that will be origin/destinations (ODs) for walking users. At different periods of each day, different sub-events will be held in different venues. During walking outdoors between ODs, pedestrians are at risk of heatstroke. In this study, many factors that affect the heatstroke risk of pedestrians, such as walking distance, solar radiation, pedestrian flow (which is represented simply by 'flow' in the following) density, the number and location of relief stations and the supply volume, are taken into the consideration.

The proposed problem is set as optimizing the number, location and supply volume of each temporary station as well as the pedestrian routes in the road network to reduce heatstroke each day during the large outdoor event. Noted that the solution of the problem corresponds to the scheduling scheme for all the sub-events every single day during the event. Before the optimization, there are some preparatory works needed to be done: (a) facilities and POIs extraction; (b) extraction and simplification of the road network for a given area; (c) the extraction and calculation of the heatstroke related data; (d) event schedule collection and pedestrian flow simulation. By handling the preparatory work and solving the optimization problem, we are not only able to provide event holders with a reasonable allocation scheme of relief stations and supply volume but also to recommend walking paths for pedestrians.

The optimization setting in our research can be described as follows:

Assumption:

(1). There will be several inflows before each individual event and outflows after the event between event venues and other facilities such as places of interests (POIs), hotels and stations.

(2). The location of the relief station cannot be changed during the day, the supply volume can be however reassigned at different times of the day.

Input:

(1). The road network information including nodes, edges, the length of each edge, and the factor value that increases the vulnerability of each edge, the coordinates of each node.

(2). The Environment related data of the given area.

(3). The numbers of time units and simulated flow density at each time interval.

(4). The sets of start nodes and end nodes of all flows.

(5). The maximum number of relief stations and supply volumes of each station, the maximum allowed heatstroke risk of all edges (road) in the road network, the edge set on the path of each flow and the edge set between every two adjacent stations on the path of each flow.

Determine:

(1). The numbers and locations of relief stations on a given day.

(2). The supply volume of each relief stations at different time units.

(3). The optimal route is made up of a set of road segments with the least heatstroke risk of different flows at different time units.

To mathematically model the problem, we represent the road network as a graph $G=(V,E)$ that consists of the vertex set V and edge set E. Specifically, a vertex $v \in V$ of the road network represents either an origin or destination point of flows or the intersections of the road segments. Then we propose a heatstroke risk metric to measure the risk of each road segment at different times during the events with different indicators including both pre-calculated parameters and the decision variables to be optimized. With the objective function of minimizing global risk values, a Mixed Integer Nonlinear Programming (MINLP) model[26] is established to work out the optimal solutions. All notations of the mathematical models that we are going to introduce are listed and described in the section of Nomenclature.

3.2 Measuring heatstroke risk

We use a framework of a traditional risk model with heterogeneous data collected from different data sources. A traditional risk model divides risk into three factors which are hazard, vulnerability and exposure. Then a simple approach for measuring emergency risk is realized by multiplication of these three factors: $R = r_{hazard} \times r_{vulnerability} \times r_{exposurre}$ .Generally, the hazard represents the possibility that the emergency happens[27] while vulnerability represents the lack of proper resistance to the emergency, which is dependent on the context information. Finally, exposure refers to the amount of time spent when exposed to the hazard or the number of people involved. Although several studies have been applied to estimate the heatstroke risk on a macro scale[28], few focus on microanalysis, which should have different indicators depending on the distinct micro context. We utilize different micro indicators to implement our risk model for micro heatstroke analysis. The indicators of the hazard, vulnerability and exposure factors are listed as follows:

3.2.1 Hazard

Hazard is measured by Wet Bulb Globe Temperature (WBGT) which has been applied in other heatstroke related studies[28, 29]. Specifically, we choose a data-driven approach to measure heatstroke hazard via historical WBGT data at different hours during summertime. We utilize a normalized index W_t to represent the probability of the time span that exceeds the warning temperature as suggested by the local government[1] on all summer days.

3.2.2 Vulnerability

Vulnerability is the factor related to the contextual environment. Generally, the vulnerability could be generated by the existing contextual environment, or reduced by improving the environment via temporary service. In this study, vulnerability is denoted by the indicators of a road segment. Sky view factor (SVF) is utilized to measure the existing contextual environment, which has been proved to be quite an important indicator for computing solar radiation related to heatstroke[30] and is very sensitive to the contextual environment. On the other hand, the relief stations are set to reduce vulnerability and a station with larger service volume (e.g., more volunteers) could help more pedestrians. Therefore, Vulnerability V could be measured by Eq. (1) with a given

[1] https://www.wbgt.env.go.jp/en/wbgt.php

SVF value V_i^I and the vulnerability reduction indicator $V_{i,t}^R$ computed from the supply volume $N_{i,t}^V$ in each road segment *i* and time interval *t*.

$$R_{i,t}^V = \frac{V_i^I}{V_{i,t}^R} = \frac{V_i^I}{1+\left(B_s^S N_{i,t}^V\right)^d} \tag{1}$$

3.2.3 Exposure

Exposure is measured by the total walking time of all pedestrians for each road segment. To simplify the calculation, we assume all pedestrians walking at the same speed in all road segments and within all time periods. As a result, the exposure of heatstroke in this study is proportional to the number of pedestrians and the road length for each road segment. Therefore, for a given road segment *i* at time interval *t*, exposure volume is denoted by Eq. (2)

$$R_{i,t}^E = L_i \sum_f B_{i,f,t}^P N_{f,t}^P \tag{2}$$

With the factors defined, the risk for flow *f* in road segment *i* at time interval *t* could be denoted by Eq. (3).

$$\begin{aligned} R_{i,f,t} &= B_{i,f,t}^P \left(W_t\right)^a \left(R_{i,t}^V\right)^b \left(R_{i,t}^E\right)^c \\ &= B_{i,f,t}^P \left(W_t\right)^a \left(\frac{V_i^I}{1+\left(B_s^S N_{i,t}^V\right)^d}\right)^b \left(L_i \sum_f B_{i,f,t}^P N_{f,t}^P\right)^c \quad \forall i \in E, f \in F, t \in T \end{aligned} \tag{3}$$

3.3 Optimization model

With the risk metric defined above, this study provides an MINLP model to work out the solutions for facility and path optimization with one objective function and several constraints. The model is solved by genetic algorithm with several strategies to accelerate the computation process.

3.3.1 Objective Function

The objective function is to minimize the total heatstroke risk, i.e., the risk value generated by the risk metric introduced in the last section, for all flows during all the events within a time period *T*, which can be denoted by Eq. (4):

$$\min \ Risk = \sum_i \sum_f \sum_t R_{i,f,t} \quad \forall i \in E, f \in F, t \in T \tag{4}$$

3.3.2 Constraints

In this study, several constraints are set either for solutions to meet the predefined parameters or for overcoming the shortages of a single objective function for enabling more practical application. Generally, constraints can be categorized into the following four groups based on their target:

A. Station constraints

The total number of supply stations established cannot exceed the maximum. It is described as below:

$$B_i^S = \begin{cases} 1, & if \ \exists i = E_s^S \\ 0, & if \ \forall i \neq E_s^S \end{cases} \quad \forall i \in E, s \in S \tag{5}$$

$$\sum_i B_i^S \leq N^{S,max} \quad \forall i \in E \tag{6}$$

B. Flow path constraints

In this study, since the flows are represented by a set of edges from the given origin and destination nodes, the path constraints mainly focus on edge connectivity and origin-destination connectivity.

Specifically, the origin-destination connectivity constraint denotes that the start (end) edge should be the only edge connected to the origin (destination) node. These constraints are as follows:

$$\sum_i B_{i,f,t}^P = 1 \quad \forall i \in E^{SE}, f \in F, t \in T \tag{7}$$

$$\sum_i B_{i,f,t}^P = 1 \quad \forall i \in E^{EE}, f \in F, t \in T \tag{8}$$

For each flow path at time interval *t*, the connectivity is judged by two Boolean matrices with the size of $n_{f,t}^E$, adjacency matrix $A_{f,t}$ and the reachability matrix $P_{f,t}$ of the selected edges. Adjacency is represented by Constraint (11) which means if there is a point connected by both edge *i* and edge *j*, then the two edges are adjacent, the value of $a_{i,j,f,t}$ is 1, otherwise it is 0. Reachability matrix $P_{f,t}$ is obtained by Boolean addition and Boolean multiplication for adjacency matrix which are described by Eq. (13) and Eq. (14), and constraint (15) ensures that the graph composed of the selected edges is connected.

$$n_{f,t}^{E}=\sum_{i}B_{i,f,t}^{P}\quad \forall i\in E,f\in F,t\in T \tag{9}$$

$$A_{f,t}=\left(a_{i,j,f,t}\right)_{n_{f,t}^{E}\times n_{f,t}^{E}}\quad \forall i,j\in E^{FP},f\in F,t\in T \tag{10}$$

$$a_{i,j,f,t}=\begin{cases}1,\ \ if\ N_{i,f,t}^{EF}=N_{j,f,t}^{EF}\ or\ N_{i,f,t}^{EF}=N_{j,f,t}^{ES}\ or\\ \quad N_{i,f,t}^{ES}=N_{j,f,t}^{EF}\ or\ N_{i,f,t}^{ES}=N_{j,f,t}^{ES} \qquad \forall i,j\in E^{FP},f\in F,t\in T\\ 0,\ \ otherwise\end{cases} \tag{11}$$

$$P_{f,t}=\left(p_{i,j,f,t}\right)_{n_{f,t}^{E}\times n_{f,t}^{E}}\quad \forall i,j\in E^{FP},f\in F,t\in T \tag{12}$$

$$P_{f,t}=\bigcup_{k=1}^{n_{f,t}^{E}}A_{f,t}{}^{(k)}\quad \forall f\in F,t\in T \tag{13}$$

$$A_{f,t}{}^{(k)}=A_{f,t}{}^{(k-1)}\ \text{e}\ A_{f,t}\quad \forall f\in F,t\in T \tag{14}$$

$$p_{i,j,f,t}=1\quad \forall i,j\in E^{FP},f\in F,t\in T \tag{15}$$

C. Volume constraints

The following volume constraints ensure that the total volume of service number of all supply stations at any time interval cannot exceed the max value. In addition, the constraint that the volume number of supply stations on each road segment should not exceed the maximum function is realized by implementing a sufficiently large constant *M*.

$$N_{i,t}^{V}=\begin{cases}\sum_{s}N_{s,t}^{V}\ ,\ \ if\ \ i=E_{s}^{S}\\ 0,\ \ otherwise\end{cases}\quad \forall i\in E,s\in S,t\in T \tag{16}$$

$$\sum_{i}\left(B_{i}^{S}N_{i,t}^{V}\right)\le N^{V,max}\quad \forall i\in E,t\in T \tag{17}$$

$$N_{i,t}^{V}\ge M\left(B_{i}^{S}-1\right)+1\quad \forall i\in E,t\in T \tag{18}$$

$$N_{i,t}^{V}\le M\left(1-B_{i}^{S}\right)+N^{SV,max}\quad \forall i\in E,t\in T \tag{19}$$

D. Risk constraint

Risk constraints are set to exclude those solutions with concentrated stations in adjacent road segments. Specifically, the constraints should ensure that the value of the risks at each edge, at all edges of each flow and between every two adjacent stations on the path of flow *f* should not exceed the predetermined max risk values. The constraints are listed as follows:

$$R_{i,t} \le Ri^{max} \quad \forall i \in E, t \in T \tag{20}$$

$$\sum_{f} R_{i,f,t} \le Rf^{max} \quad \forall i \in E, f \in F, t \in T \tag{21}$$

$$R_{s,k,f,t} \le Rs^{max} \quad \forall s,k \in S, f \in F, t \in T \tag{22}$$

3.4 Optimization algorithm

MINLP problem is NP-hard and it is time-consuming to directly apply any solution to the proposed model due to the large number of variables. We then apply several strategies to effectively generate efficient solutions.

In particular, the problem is solved by genetic algorithm (GA). GA is a widely used effective algorithm with good performance of global search and strong robustness. It is suitable for solving complex optimization problems that can be described as mixed linear models or mixed nonlinear models. The GA algorithm in this work starts with a set of the initial population that is generated based on certain requirements rather than randomly generated. This is beneficial to improve the convergence rate and the quality of the solution. Then the algorithm evaluates each individual in the population through the fitness function. A certain proportion of individuals at the top are selected as elite individuals and directly retained in the next-generation population. The next-generation population also includes children who are reproduced by selected elite individuals through crossover and mutation. With the process of evolution, the solution gradually approaches the optimal solution according to the principle of the survival of the fittest.

3.4.1 Generating initial GA population

Inspired by the assignment of initial solutions to ant colony optimization (ACO) in[31], the proposed method generated an initial GA population with feasible solutions to accelerate the convergence. In this study, we choose a group of paths based on the actual situation of excluding the solutions with large detours. Specifically, we apply Dijkstra algorithm to generate the shortest paths with the least SVF weighted length. Then a breadth-first search is applied to acquire all paths with the edge number smaller than or equal to the edge number in the shortest path + 5. Then the initial population can be represented by a combination of different paths for different flows.

3.4.2 Penalty function

To get solutions that satisfy constraints in the presented model, we use a penalty function so as to exclude the chromosomes that cannot meet the constraints in the evolution. Then the fitness function by which the next generation is bred could be represented by a sum of the objective function and penalty function, which is shown as Eq. (23).

$$F(\mathbf{X},\delta_n) = f(\mathbf{X}) + \delta_1 \sum_{i=1}^{n^p} \left(p_i(\mathbf{X}) - 1 \right) + \delta_2 \sum_{j=1}^{n^v} \max\left\{0, \left(v_i(\mathbf{X}) - N^{V,max} \right)\right\} + \delta_3 \sum_{k=1}^{n^{ri}} \max\left\{0, \left(ri_i(\mathbf{X}) - Ri^{max} \right)\right\} + \delta_4 \sum_{p=1}^{n^{rf}} \max\left\{0, \left(rf_i(\mathbf{X}) - Rf^{max} \right)\right\} + \delta_5 \sum_{q=1}^{n^{rs}} \max\left\{0, \left(rs_i(\mathbf{X}) - Rs^{max} \right)\right\} \tag{23}$$

Where **X** represents a chromosome, i.e., a potential solution to the problem. $f(\mathbf{X})$ is the objective function while the remaining items are penalty functions. $\delta_k \left(k = 1,2,3,4,5\right)$ represents the weight of each penalty function which is usually a sufficiently large constant. The more variables that do not meet the constraints, the greater the value of the penalty function. Besides, the objective function is to minimize the risk of heatstroke. Therefore, the smaller the value of fitness, the better the solution. In addition, to further speed up the convergence, we compute the fitness function of the entire population in parallel.

4. Case study: an application scenario for Tokyo Olympic Games in Tokyo Waterfront City

4.1 Study area

Based on the model proposed above we conduct experiments on the walkable space of Olympic venues in the Tokyo Waterfront City (TWC) which mainly includes the regions of Odaiba (Aomi included) and Ariake. During the Tokyo Olympic Games, a lot of games are scheduled to be held in TWC. Besides, as a region with concentrated scenic spots, shopping malls and theme parks, TWC attracts a lot of visitors every year and ranks 12th among 4,027 scenic spots in the central Tokyo

area[2]. The abundant scenic spots and hotels distributed in TWC make it a space with the forecasted high demand for walking during the Olympic Games as there will be a lot of visitors walking to the scenic spots near Olympic venues[32]. Figure 1 shows the map of the two main regions in TWC with different types of facilities and the extracted road network.

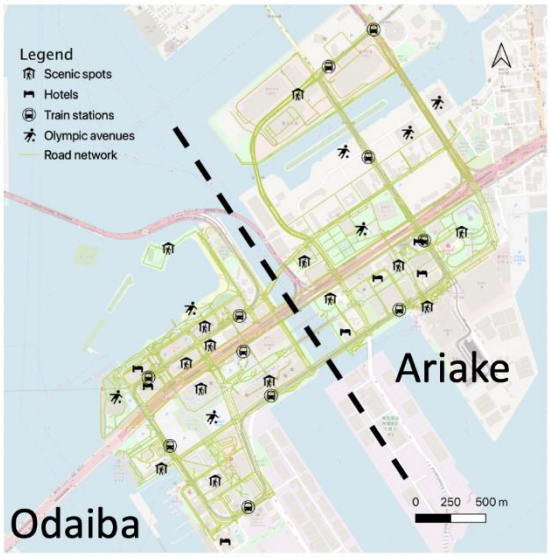


Fig. 1 Spatial distribution of POIs in TWC

4.2 Data collection and preprocessing

4.2.1 POI and facility extraction

We collect different types of POIs including scenic spots, hotels, railway stations and Olympic venues from heterogeneous data sources. Scenic spots and hotel data are taken from TripAdvisor, railway stations from the National Land Numerical Information and avenue locations are taken from their official websites. Starting with the initial number of 400 raw POIs in Odaiba area collected from the TripAdvisor, we next merge them based on their spatial entities to remove duplicated POIs in the same building and we use the total comment numbers in each location as its popularity.

4.2.2 Road network

Road network is collected from OpenStreetMap (OSM) with the help of the library Osmnx [33]. In TWC region, the raw data collected from OSM include more than 4,000 road links including different road types, which makes it difficult to be directly applied to the optimization problem for

[2] https://www.tripadvisor.jp/Attractions-g298184-Activities-a_allAttractions.true-Tokyo_Tokyo_Prefecture_Kanto.html, based on the ranking on Aug 25, 2020

pedestrians. Thus, we simplify the road network based on the extracted skeleton[34, 35] and attributes of roads: levels and types. After this simplification, the total number of road segments in TWC area is reduced to 234, in which 131 segments are located in Odaiba and 103 segments are located in Ariake.

4.2.3 Heatstroke related data collection and processing

A. WBGT data

WBGT data is collected from the government website[3] at one-hour interval. To represent the situation during the Tokyo Olympic Games, we take data on the same day from 2017 to 2019 to calculate the normalized index.

B. SVF data

In order to calculate SVF for each road segment, we refer to the work of[35] to collect Google Street View images for each simplified road segment in the TWC area and we conduct image segmentation to extract the sky range in the images. Specifically, we generate intermediate points on each road segment with a five-meter interval, then use the coordinates of the points and nodes as the request parameters for Google API to gather panorama data. Having collected the panorama data, we utilize a SegNet model[36] trained by CityScape dataset[37] for image detection and convert the detection results to fisheye images to calculate SVF values.

4.2.4 Event schedule collection and flow simulation

In the case study, events represent the sports held in the venues of TWC during the Olympic Games. For each sport event, the official schedule of time, location[4] and the estimated audience number[5] are collected from the official website. In total there will be 71 sports events held during 16 days in the whole research area.

To simplify the computation, we choose in this study an hour as the time unit for calculating the flows. For each event, its estimated audience number is distributed as the total inflows within

[3] https://www.wbgt.env.go.jp/wbgt_data.php

[4] https://tokyo2020.org/ja/schedule, in this study we use the schedule before the games' postponement.

[5] https://www.shochi-honbu.metro.tokyo.lg.jp/TOKYO2016_15_9.pdf, in this study we use the schedule before the games' postponement.

two-time units (two hours) just before the event. Similarly, the audience number is taken to represent the total outflows distributed during the two-time units (two hours) right after the event. Since there are no records for allocating the total flow volumes to individual flows, we apply the Huff model[38] to simulate the flow number between venues and other facilities which takes both distance decay and the facility popularity into consideration.

4.3 Experiments

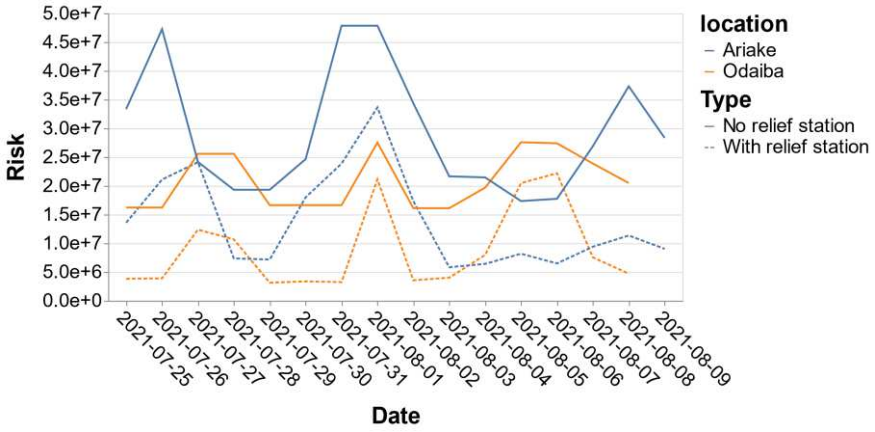
Having the data collected and processed as explained above, we conduct in this case study several experiments with different data inputs for solving the optimization problem under different flow numbers and contextual information.

On one hand, TWC has an area of 400ha which is too large to form a single walkable space and the distance of facilities between its two regions (Odaiba and Ariake) is relatively far (as shown in Figure 1; it is necessary to go across bridges to reach another region). Thus, in this study we regard these two regions as two independent walkable spaces and conduct experiments separately on each of these two regions. On the other hand, the sports events on different days have different schedules and different estimated number of audiences, while there is sufficient time to shift temporary stations and supplies at night. Therefore, in this study we build different models for different days to make the application scenario more practical and to reduce the computation time for each model. In total, there will be 32 models (a combination of 16 days and 2 walkable spaces) to solve the problem in different contextual environments and people flows with a single group of parameters.

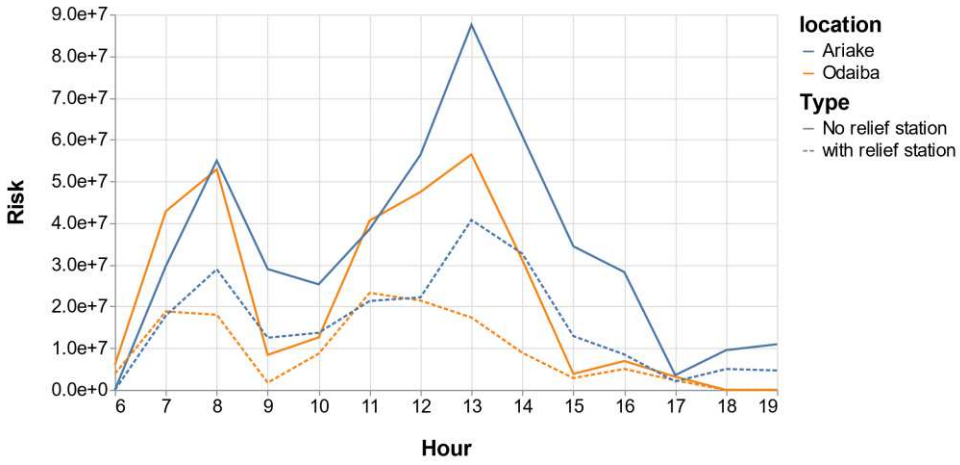
The optimization is implemented by C++ in Ubuntu Linux 14.04.2. There are 2 CPUs where each of which is Intel (R) Xeon (R) CPU E5-2699 v3 @2.30 GHz. For the parameters of GA, the generation is set as 2000, the population size is set as 3000, while the rates of selection, crossover and mutation are 0.4, 0.8 and 0.3, respectively.

4.3.1 Result statistic and visualization

The results and the visualizations use the solution with a group of parameters in which the station number is 10, the total supply volume is 100 and the maximum supply volume in each station is set to 20. The two plots in Figure 2 respectively show the optimized total risk as well as the expected risk for the shortest paths without relief stations in different days and hours in different areas during Olympic games.



(a) Daily risk



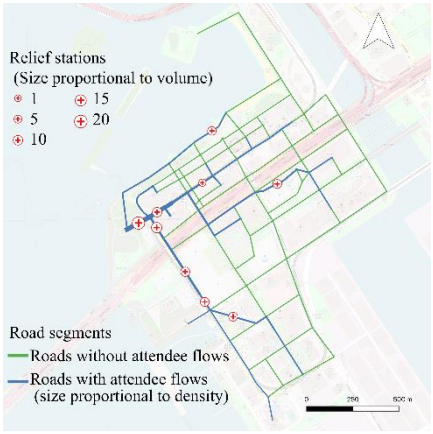
(b) Hourly risk

Fig.2 Risk in Odaiba and Ariake area during Olympic Games

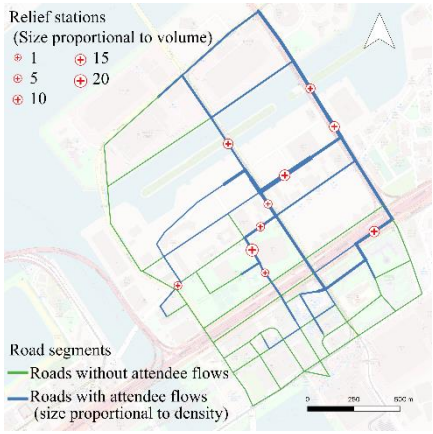
From the plots, we can find that the risk varies in different days and hours with several peaks in several days and hours. Daily and hourly differences indicate that heatstroke risk is very sensitive to the Olympic schedule. Additionally, hourly differences can also reflect weather variation within a day. This is noticeable in the daily risk change, besides the peak observed at 1 p.m., due to both the busy schedule of events and high temperature. The observed peak at 8 a.m. is mainly due to the game schedule. This result mainly could contribute to the efforts made by the Tokyo government

aiming at reducing the heatstroke risk for outdoor sports events. Although the performance varies at different hours and days, correctly setting relief stations and optimizing routes can significantly reduce total risk in different event scenarios.

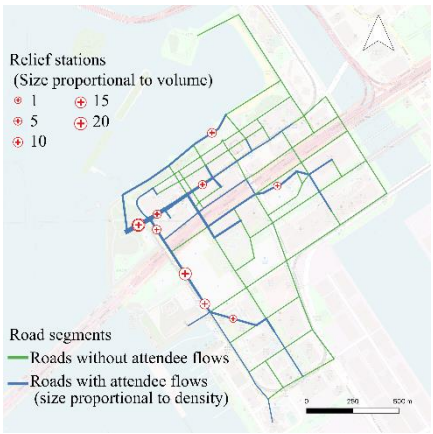
Since there are too many events to be visualized in maps, we selectively visualize the flow density, optimized facility location and the supply volume of each facility in both Odaiba and Ariake area at 8 a.m. and 1 p.m. on July 26 and Aug 1 respectively in 8 maps of Figure 3. From these maps, we can observe that the supply volume and pedestrian flow density varies at different hours in a day, which stresses the significance of optimizing supply volume within one day. On the other hand, the changes in the optimized locations and flow density at different days suggests the necessity of setting different relief stations on different days.



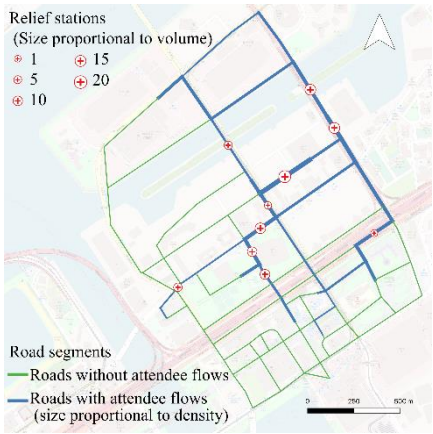
(a) Odaiba at 8 a.m., July 26



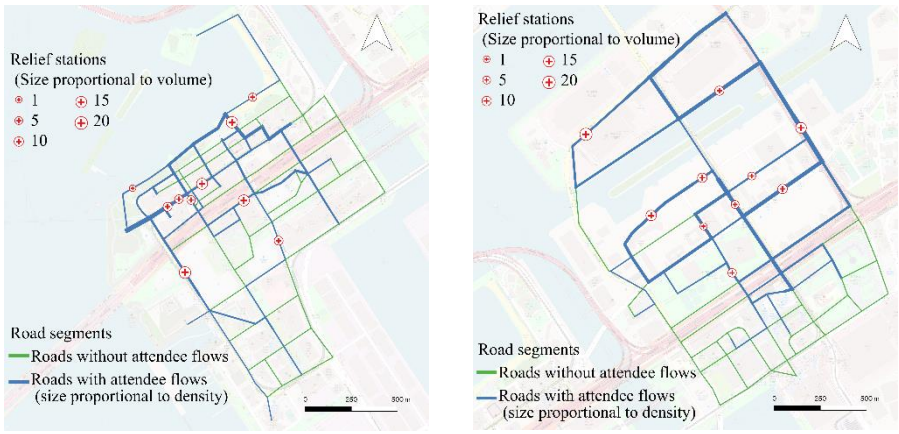
(b) Ariake at 8 a.m., July 26



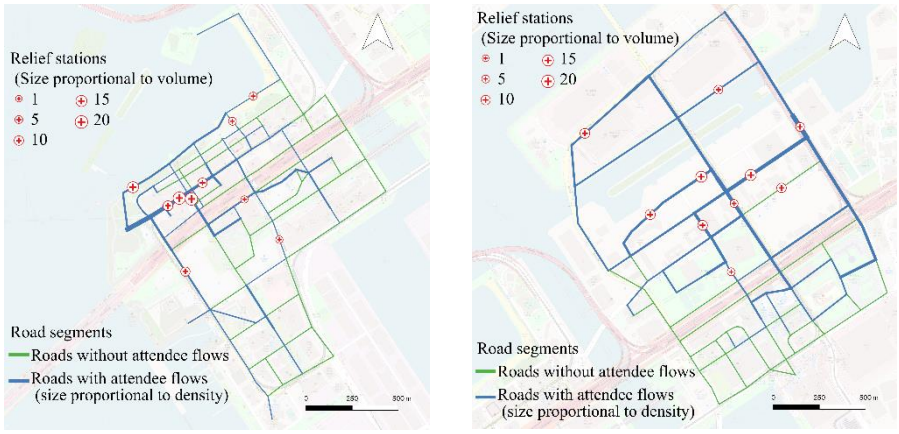
(c) Odaiba at 1 p.m., July 26



(d) Ariake at 1 p.m., July 26



(e) Odaiba at 8 a.m., Aug 1 (f) Ariake at 8 a.m., Aug 1

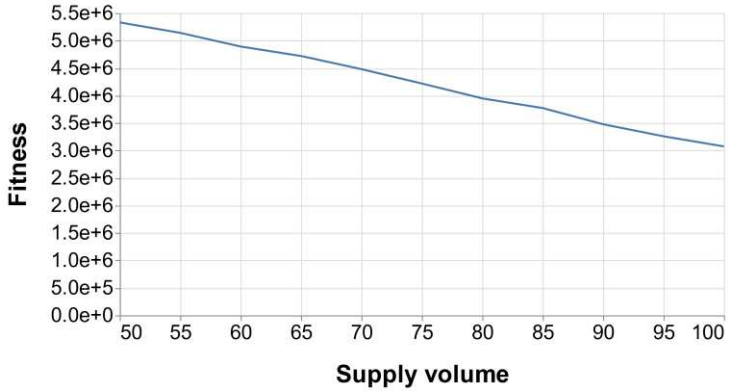


(g) Odaiba at 1 p.m., Aug 1 (h) Ariake at 1 p.m., Aug 1

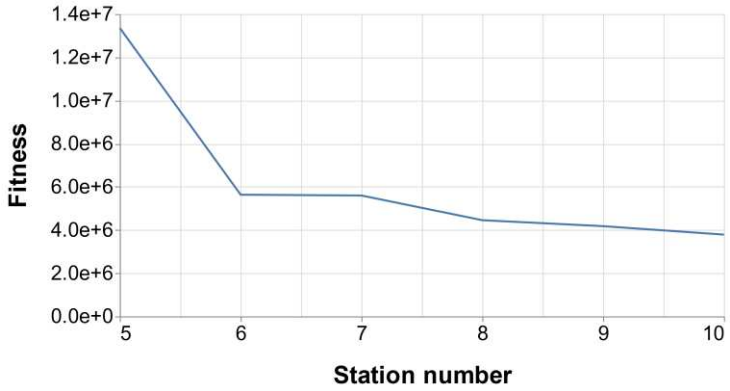
Fig. 3 Visualization of event flow density, relief station location and supply volume in each station

4.3.2 Sensitivity analysis

We now conduct sensitivity analysis on the supply volume and stations. Figure 4 shows the sensitivity results for the data of Odaiba on the 3rd, Aug.



(a) Sensitivity analysis of total supply volume in Odaiba, Aug 3



(b) Sensitivity analysis of station number in Odaiba, Aug 3

Fig. 4 Sensitivity analysis: Fitness functions in relation to station numbers and supply volumes

The increase in the total number of supply volume or in the number of relief stations will reduce the value of the fitness function. This means that increasing the number of supplies (e.g., the number of volunteers or the amount of relief resources such as bottled water) or relief stations will either reduce the total risk or reduce the probability that the variables will not meet the constraints in the presented model.

From the results shown in Figure 4(a), the fitness basically decreases linearly with the increase in the number of supply volume. This shows that in the experiments, 100 units of supply volume may not have reached the upper limit to minimize the total risk in Odaiba area on 3rd, Aug. However, for the number of relief stations, as shown in Figure 4(b), the rate of fitness reduction gradually decreases along with the increase in the number of stations and tends to remain stable.

On the other hand, when the number of supply volume and the number of relief stations are greater than 50 and 5, respectively, for each flow, the risk between adjacent relief stations on its path has already met the constraint of being less than the maximum risk. Therefore, these solutions are feasible. Overall, when resources are limited, such as the maximum supply volume is 100 and the maximum number of relief stations is 10, the more supply volumes and relief stations, the lower the total risk value is.

4.3.3 Ablation study

We conduct several ablation analyses to directly evaluate the performance of our method. Specifically, we compare the fitness between our model and other ablated model settings which are listed as follows:

(1). Fixed routes: with the optimized station location, we set all routes of each flow as the shortest routes.

(2). Fixed volume: with the optimized station location and routes, we set fixed and equal supply volume, i.e., 10 volume units in each station when the total volume is 100 and the station number is 10.

(3). No station: only routes are optimized for each OD and no relief stations are set.

Table 1 Ablation analysis based on fitness

	Odaiba				Ariake			
Date 2021	Our method	Fixed routes	Fixed volume	No station	Our method	Fixed routes	Fixed volume	No station
07-25	**3.83E+06**	3.99E+06	3.87E+06	1.62E+07	**1.36E+07**	**1.36E+07**	1.39E+07	3.33E+07
07-26	**3.90E+06**	4.31E+06	3.96E+06	1.62E+07	**3.11E+07**	**3.11E+07**	4.11E+07	7.73E+07
07-27	**3.24E+07**	3.27E+07	7.27E+07	4.56E+07	**5.41E+07**	5.82E+07	5.43E+07	6.42E+07
07-28	**4.07E+07**	4.08E+07	4.08E+07	6.56E+07	**7.37E+06**	1.71E+07	7.49E+06	3.93E+07
07-29	**3.15E+06**	3.61E+06	3.26E+06	1.66E+07	**7.18E+06**	7.24E+06	7.29E+06	2.93E+07
07-30	**3.38E+06**	2.35E+07	3.53E+06	3.66E+07	**5.80E+07**	6.68E+07	5.80E+07	7.46E+07
07-31	**3.26E+06**	3.40E+06	1.33E+07	3.66E+07	**4.39E+07**	5.37E+07	4.40E+07	7.79E+07
08-01	**3.11E+07**	3.15E+07	3.12E+07	5.75E+07	**8.36E+07**	9.62E+07	8.38E+07	1.18E+08
08-02	**3.57E+06**	3.76E+06	3.67E+06	5.61E+07	**2.73E+07**	5.62E+07	2.74E+07	7.44E+07
08-03	**4.01E+06**	4.26E+06	4.10E+06	3.61E+07	**5.84E+06**	5.88E+06	5.94E+06	2.17E+07
08-04	**1.80E+07**	3.80E+07	1.81E+07	4.97E+07	**6.42E+06**	**6.42E+06**	6.53E+06	2.15E+07
08-05	**5.05E+07**	6.06E+07	5.05E+07	7.76E+07	**8.19E+06**	2.81E+07	8.23E+06	3.73E+07
08-06	**4.22E+07**	6.28E+07	4.23E+07	6.74E+07	**2.65E+07**	3.65E+07	2.65E+07	4.77E+07
08-07	**1.76E+07**	2.76E+07	1.77E+07	4.39E+07	**9.40E+06**	9.57E+06	9.68E+06	4.69E+07
08-08	**4.72E+06**	4.78E+06	4.99E+06	2.05E+07	**1.14E+07**	2.14E+07	1.14E+07	5.73E+07
08-09					**9.04E+06**	1.90E+07	9.17E+06	4.83E+07

The ablation analysis results conducted on different dates and areas are given in Table 1. Generally, we can observe an increase of the fitness value under different tested assumptions, which results from poor performance of stations and routes under the risk model. Also, we can observe some large difference between the ablated results and our models, which results from the large penalty value generated in the ablated models.

5. Conclusion

This study proposes a framework to evaluate the heatstroke risk in walkable spaces during large events by utilizing context-aware indicators that are generated by large and heterogeneous data including facilities, road networks and street view images. We propose to solve an MINLP problem for optimizing routes of pedestrians, determining the location of relief stations and the supply volume in each relief station. We then conduct a case study on the planned site of the Tokyo Olympics that is simulated during the two weeks' long period of the Olympic schedule.

With the indicators and metrics for other types of risks, our method is potentially applicable for solving optimization problems in other specific scenarios that require facility and route optimization in large walkable spaces such as large theme parks and big outdoor exhibitions.

This study has the following limitations on data and models. First, the population simulation could not be evaluated in the current stage since there was no such a big event held in TWC before, as well as it is difficult to find other events with flow data and detailed schedules being provided. Besides, although various prior studies support our choice of contextual information for modeling the heatstroke risk, it is still an assumption that the contextual information utilized in this study has an actually high influence in a specific scenario that we focus on. Besides, we use MINLP model for the optimization with tens of thousands of variables, which suggests that it needs more validation for the cases of larger and more complicated walkable spaces than ones used for the Tokyo Olympics.

In the future study, we will first try to improve the limitations of our work. Specifically, we will simulate pedestrian flows and evaluate contextual information with more observational data from several datasets. Besides we will propose better models that apply improved, more effective algorithms to optimize routes and facilities. Finally, with the improved data and methods, we will try to apply our framework to other application scenarios.

Acknowledgments

(1). The support provided by the China Scholarship Council (CSC) during a visit of Yan Wu to the University of Tokyo is gratefully acknowledged.

(2). Financial support from 19K15260 KAKENHI (Early-Career Scientists Grant) is acknowledged.

References

[1] Kwan M-P. The Uncertain Geographic Context Problem. Annals of the Association of American Geographers. 2012;102:958-68.

[2] Church R, ReVelle C. The maximal covering location problem. Papers of the regional science association: Springer-Verlag; 1974. p. 101-18.

[3] Toregas C, Swain R, ReVelle C, Bergman L. The location of emergency service facilities. Operations research. 1971;19:1363-73.

[4] Ahmadi-Javid A, Seyedi P, Syam SS. A survey of healthcare facility location. Computers & Operations Research. 2017;79:223-63.

[5] Siddiq AA, Brooks SC, Chan TC. Modeling the impact of public access defibrillator range on public location cardiac arrest coverage. Resuscitation. 2013;84:904-9.

[6] Hogan K, ReVelle C. Concepts and applications of backup coverage. Management science. 1986;32:1434-44.

[7] Larson RC. Decision models for emergency response planning. Handbook of Homeland Security: Citeseer; 2005.

[8] Schempp T, Zhang H, Schmidt A, Hong M, Akerkar R. A framework to integrate social media and authoritative data for disaster relief detection and distribution optimization. International Journal of Disaster Risk Reduction. 2019;39:101143.

[9] Oran A, Tan KC, Ooi BH, Sim M, Jaillet P. Location and routing models for emergency response plans with priorities. Future Security Research Conference: Springer; 2012. p. 129-40.

[10] Başar A, Çatay B, Ünlüyurt T. A taxonomy for emergency service station location problem. Optimization letters. 2012;6:1147-60.

[11] Daskin MS. A maximum expected covering location model: formulation, properties and heuristic solution. Transportation science. 1983;17:48-70.

[12] Coskun N, Erol R. An optimization model for locating and sizing emergency medical service stations. Journal of medical systems. 2010;34:43-9.

[13] Brown PJ, Bovey JD, Chen X. Context-aware applications: from the laboratory to the marketplace. IEEE personal communications. 1997;4:58-64.

[14] Subbu KP, Vasilakos AV. Big data for context aware computing–perspectives and challenges. Big Data Research. 2017;10:33-43.

[15] Yao L, Sheng QZ, Qin Y, Wang X, Shemshadi A, He Q. Context-aware point-of-interest recommendation using tensor factorization with social regularization. Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval2015. p. 1007-10.

[16] Chen J, Jiang W. Context-Aware Personalized POI Sequence Recommendation. International Conference on Smart City and Informatization: Springer; 2019. p. 197-210.

[17] Laß C, Herzog D, Wörndl W. Context-aware tourist trip recommendations. Proceedings of the 2nd Workshop on Recommenders in Tourism co-located with 11th ACM Conference on Recommender Systems (RecSys 2017), Como, Italy, August 272017.

[18] Siriaraya P, Wang Y, Zhang Y, Wakamiya S, Jeszenszky P, Kawai Y, et al. Beyond the Shortest Route: A Survey on Quality-Aware Route Navigation for Pedestrians. IEEE Access. 2020;8:135569-90.

[19] Gavalas D, Kasapakis V, Konstantopoulos C, Pantziou G, Vathis N.. Personal and Ubiquitous Computing. 2017;21:137-55.

[20] Zhang Y, Siriaraya P, Wang Y, Wakamiya S, Kawai Y, Jatowt A. Walking down a different path: route recommendation based on visual and facility based diversity. Companion Proceedings of the The Web Conference 20182018. p. 171-4.

[21] Mata F, Torres-Ruiz M, Guzmán G, Quintero R, Zagal-Flores R, Moreno-Ibarra M, et al. A mobile information system based on crowd-sensed and official crime data for finding safe routes: A case study of Mexico City. Mobile Information Systems. 2016;2016.

[22] Zhang Y, Siriaraya P, Kawai Y, Jatowt A. Rehab-path: Recommending alcohol and drug-free routes. Proceedings of the 28th ACM International Conference on Information and Knowledge Management2019. p. 2929-32.

[23] Bao S, Nitta T, Ishikawa K, Yanagisawa M, Togawa N. A safe and comprehensive route finding method for pedestrian based on lighting and landmark. 2016 IEEE 5th Global Conference on Consumer Electronics: IEEE; 2016. p. 1-5.

[24] Galbrun E, Pelechrinis K, Terzi E. Urban navigation beyond shortest route: The case of safe paths. Information Systems. 2016;57:160-71.

[25] Posti M, Schöning J, Häkkilä J. Unexpected journeys with the HOBBIT: the design and evaluation of an asocial hiking app. Proceedings of the 2014 conference on Designing interactive systems2014. p. 637-46.
[26] Lee J, Leyffer S. Mixed integer nonlinear programming: Springer Science & Business Media; 2011.
[27] Kron W. Keynote lecture: Flood risk= hazard× exposure× vulnerability. Flood defence. 2002:82-97.
[28] Kasai M, Okaze T, Yamamoto M, Mochida A, Hanaoka K. Summer heatstroke risk prediction for Tokyo in the 2030s based on mesoscale simulations by WRF. Journal of Heat Island Institute International Vol. 2017;12:2.
[29] Lemke B, Kjellstrom T. Calculating workplace WBGT from meteorological data: a tool for climate change assessment. Industrial Health. 2012;50:267-78.
[30] MASOUD B, COCH H, CRESPO I, BECKERS B. Effects of urban morphology on shading for Pedestrians. PLEA 2018: Smart and Healthy Within the Two-Degree Limit. 2018:1029.
[31] Zhang H, Liang Y, Liao Q, Wu M, Yan X. A hybrid computational approach for detailed scheduling of products in a pipeline with multiple pump stations. Energy. 2017;119:612-28.
[32] Weed M. Olympic tourism: Routledge; 2007.
[33] Boeing G. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Computers, Environment and Urban Systems. 2017;65:126-39.
[34] Ladak A, Martinez RB. Automated derivation of high accuracy road centrelines Thiessen polygons technique. Link: http://www esri com/library/userconf/proc96/TO400/PAP370 P. 1996;370.
[35] Liang J, Gong J, Zhang J, Li Y, Wu D, Zhang G. GSV2SVF-an interactive GIS tool for sky, tree and building view factor estimation from street view photographs. Building and Environment. 2020;168:106475.
[36] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence. 2017;39:2481-95.
[37] Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, et al. The cityscapes dataset for semantic urban scene understanding. Proceedings of the IEEE conference on computer vision and pattern recognition2016. p. 3213-23.
[38] Huff DL. Defining and estimating a trading area. Journal of marketing. 1964;28:34-8.

Nomenclature

List of abbreviations

ACO	Ant Colony Optimization
AED	Automated External Defibrillator
EFLO	Emergency Facility Location Optimization
GA	Genetic Algorithm
LBS	Location-based Services
LBSN	Location-based Social Networking
MCLP	Problem to Maximize Covering Location
MEXCLP	Maximum Expected Coverage Location Problem
MINLP	Mixed Integer Nonlinear Programming
OD	Origins and Destinations
OPAD	Off-site Public Access Devices
OSM	OpenStreetMap
POI	Points of Interest
SVF	Sky View Factor
TWC	Tokyo Waterfront City
UGCoP	Uncertain Geographic Context Problem
WBGT	Wet Bulb Globe Temperature

Indices and Sets

E	The ID set of all edges in the road network, each one denoted by indices i and j indicating start and end node of an edge
E^{EE}	The set of the edges connected with end node of each flow at all time intervals, each denoted by indices i and j, $E^{EE} \in E$

E^{FP}	The ID set of selected edges of the flow path in the road network, denoted by indices i and j, $E^{FP} \in E$
E^{SE}	The ID set of the edges connected with start node of each flow at all time intervals, each denoted by indices i and j, $E^{SE} \in E$
F	The ID set of flows, denoted by the index f
S	The ID set of stations, denoted by s and k
T	The ID set of candidate time intervals, denoted by the index t
V	The ID set of all nodes in the road network, denoted by indices α and β

Input Parameters

c	Factor value that increases vulnerability in the road segment, the average SVF value in this paper
L_i	Road length (or travel time) of edge i
N^T	The number of time intervals
N_t^F	The number of flows at time interval t
$N_{f,t}^P$	The number of people of flow f during time interval t
$N^{V,max}$	The max supply volumes of all stations
$N^{SV,max}$	The max supply volumes of one station
$N^{S,max}$	The max station numbers
$N_{i,f,t}^{EF}$	The first node ID of the selected edge i on flow path f at a time interval t
$N_{i,f,t}^{ES}$	The second node ID of the selected edge i on flow path f at a time interval t
Rf^{max}	The threshold of the maximum risk of edges of flow f
Ri^{max}	The threshold of the maximum risk of each edge
Rs^{max}	The threshold of the maximum risk between two adjacent stations on the path of flow f
W_t	Hazard factor denoted by WBGT score during a time interval

Decision variables

$B_{i,f,t}^P$	Binary variable, 1 if a flow f is observed on edge i at a time interval t, 0 otherwise
E_s^S	Integer variable refers to the ID of the edge which has a station s
$N_{s,t}^V$	Integer variable refers to the volume of service station s at a time interval t

Intermediate variables

$a_{i,j,f,t}$	The value of each element in the adjacency matrix of selected edges for flow f at time interval t
B_i^S	Binary variable, 1 if there are some stations are established on edge i; 0 otherwise
$N_{i,t}^V$	Integer variable refers to the volume of service in a station on edge i at a time interval t
$n_{f,t}^E$	The number of edges on flow path f at time interval t
$p_{i,j,f,t}$	The value of each element in the accessibility matrix of selected edges for flow f at time interval t
$R_{i,f,t}$	Refers to the risk value of edge i of flow f at a time interval t
$R_{i,t}^E$	Refers to vulnerability on edge i at a time interval t
$R_{i,t}^V$	Refers to exposure on edge i at a time interval t
$V_{i,t}^R$	Factor value that decreases vulnerability in the road segment i at time interval t

Declarations

Ethics approval and consent to participate (Not applicable)

Consent for publication (Not applicable)

Availability of data and materials

- https://www.wbgt.env.go.jp/en/wbgt.php
- https://www.tripadvisor.jp/Attractions-g298184-Activities-a_allAttractions.true-Tokyo_Tokyo_Prefecture_Kanto.html, based on the ranking on Aug 25, 2020
- https://www.wbgt.env.go.jp/wbgt_data.php
- https://tokyo2020.org/ja/schedule, in this study we use the schedule before the games' postponement.
- https://www.shochi-honbu.metro.tokyo.lg.jp/TOKYO2016_15_9.pdf, in this study we use the schedule before the games' postponement.

Competing interests

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

Funding

(1) The funding provided by China Scholarship Council (CSC)

(2) 19K15260 KAKENHI (Early-Career Scientists Grant)

Authors' contributions

- YW performed the mathematical modeling, the experiment and wrote the manuscript;
- TX contributed to the conception of the study, data processing, the visualization and was a major contributor in writing the manuscript;
- AJ and KK contributed to the manuscript revision;
- HZ helped improve the method of the optimization;
- XF and RS helped improve the quality of the article and review the manuscript.

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Figures

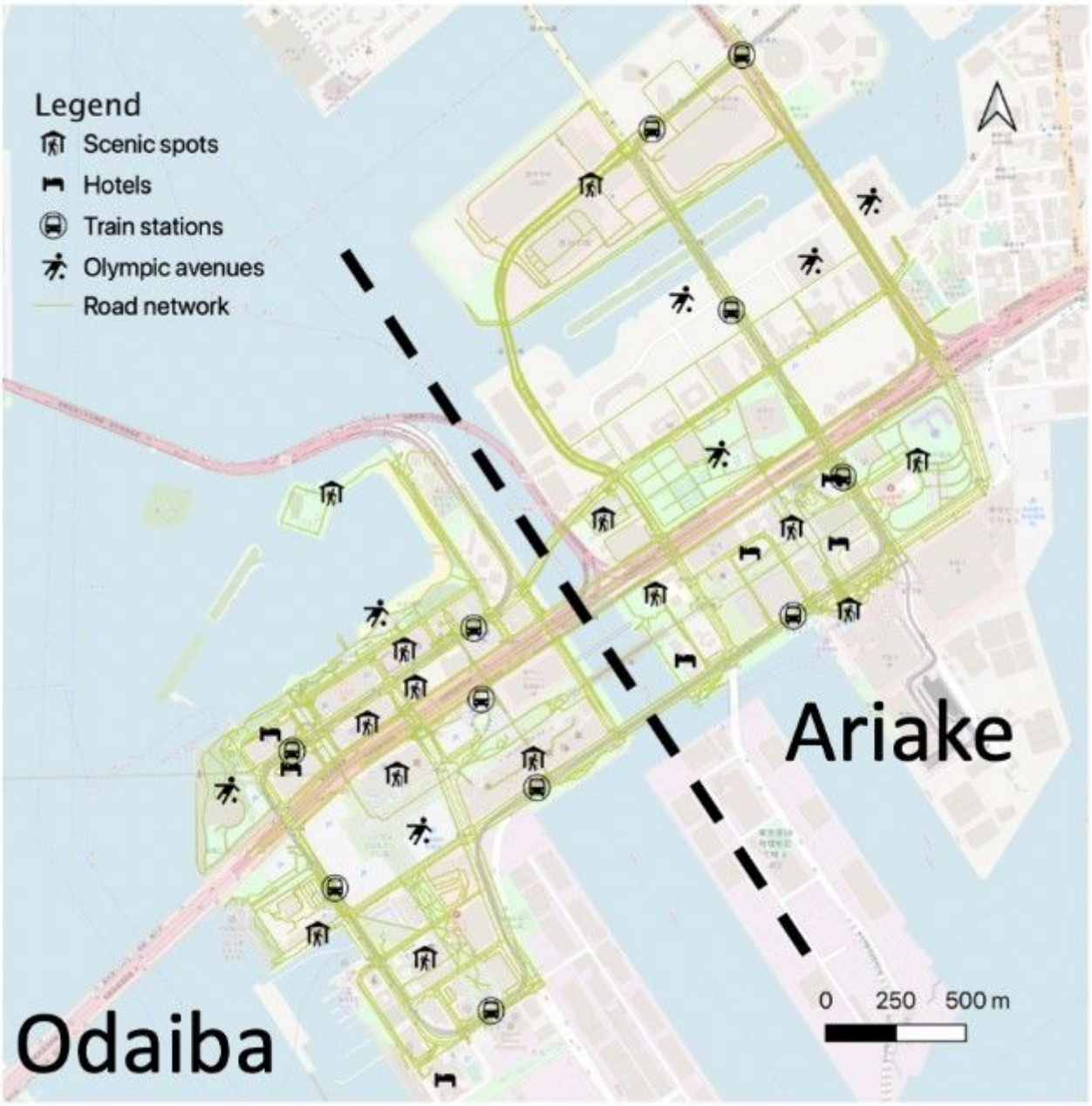
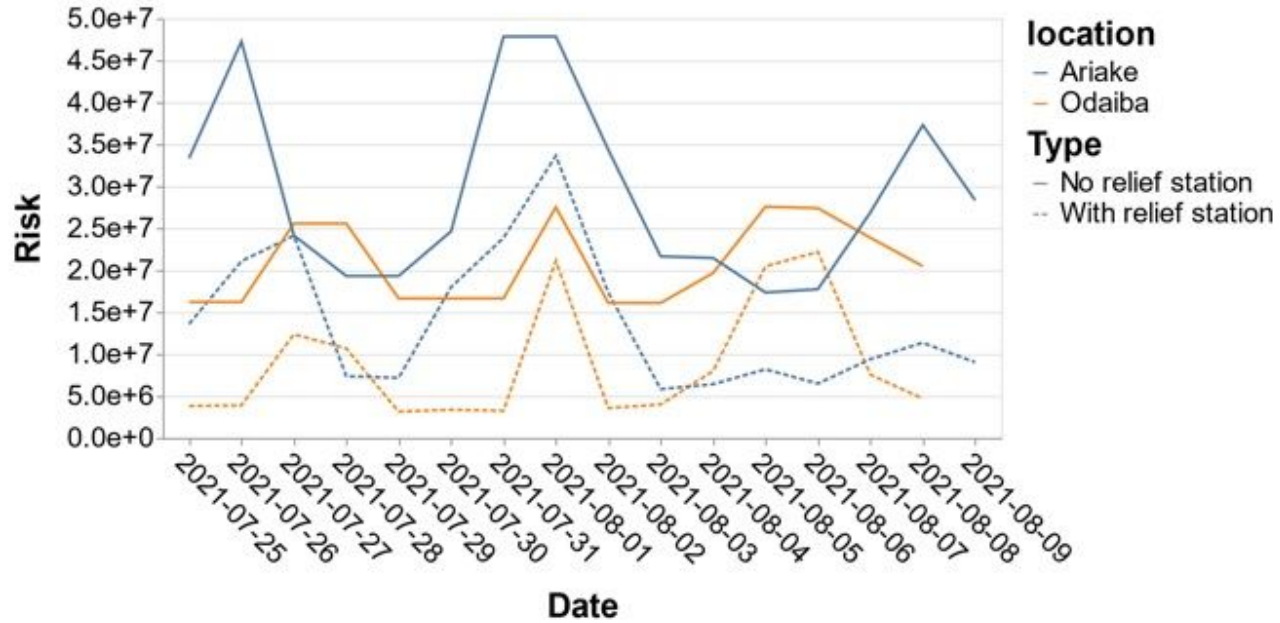
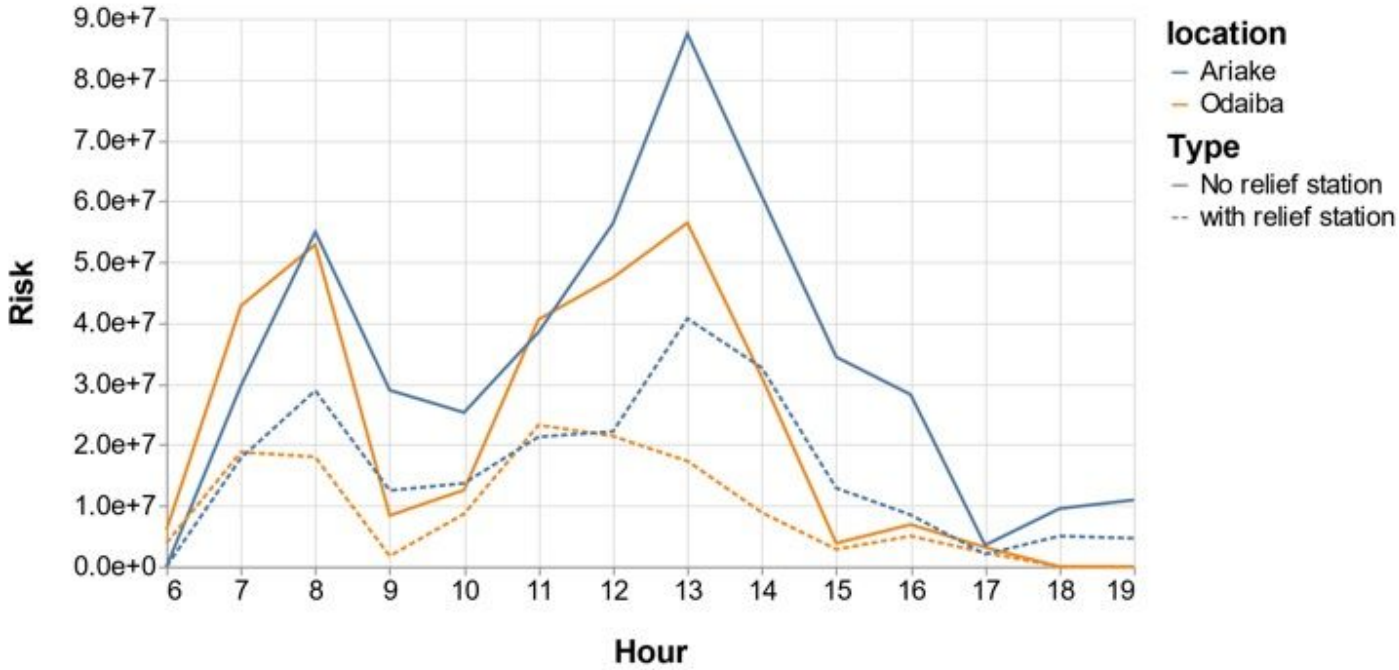


Figure 1

Spatial distribution of POIs in TWC Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.



(a) Daily risk



(b) Hourly risk

Figure 2

Risk in Odaiba and Ariake area during Olympic Games

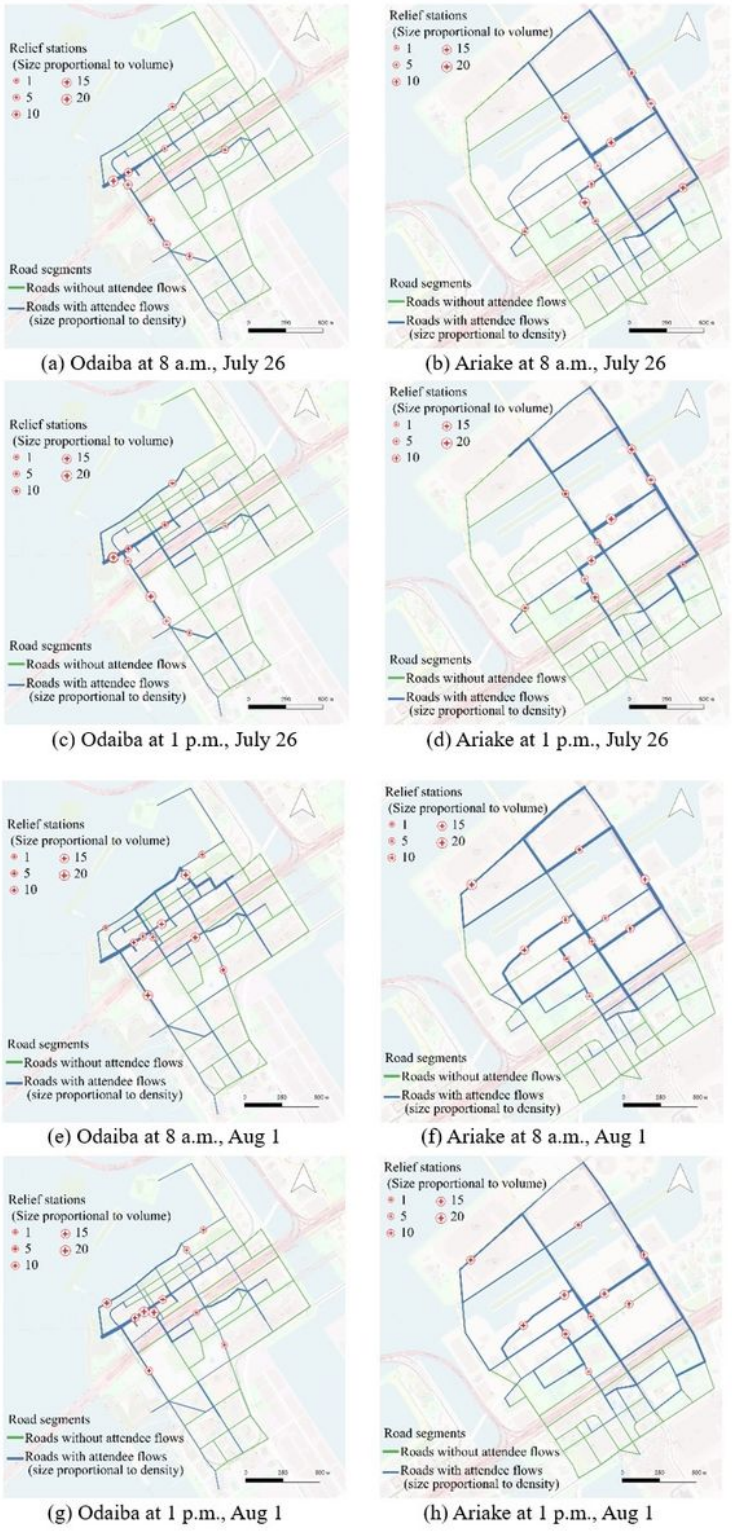
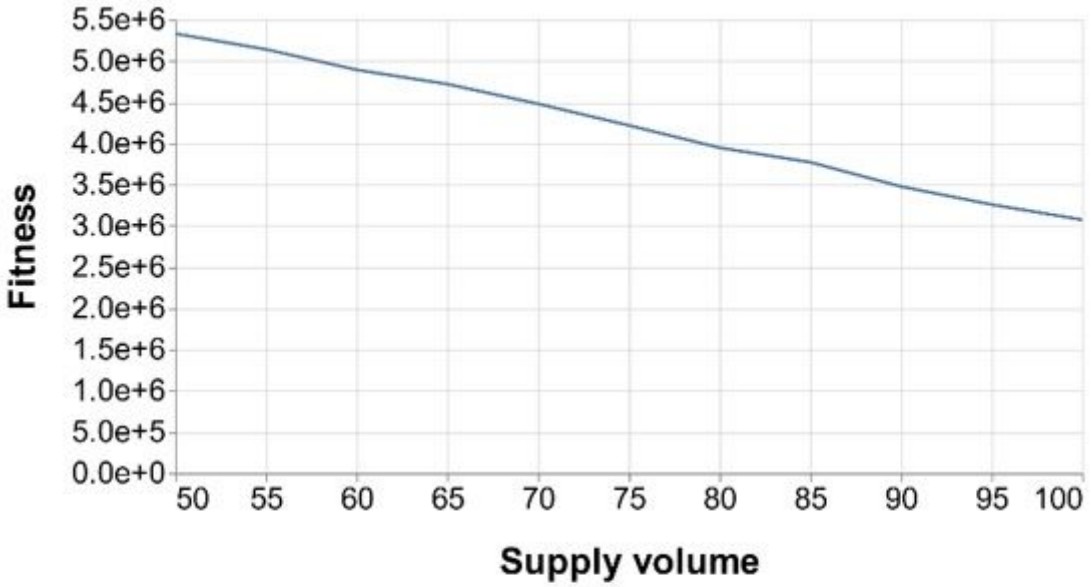
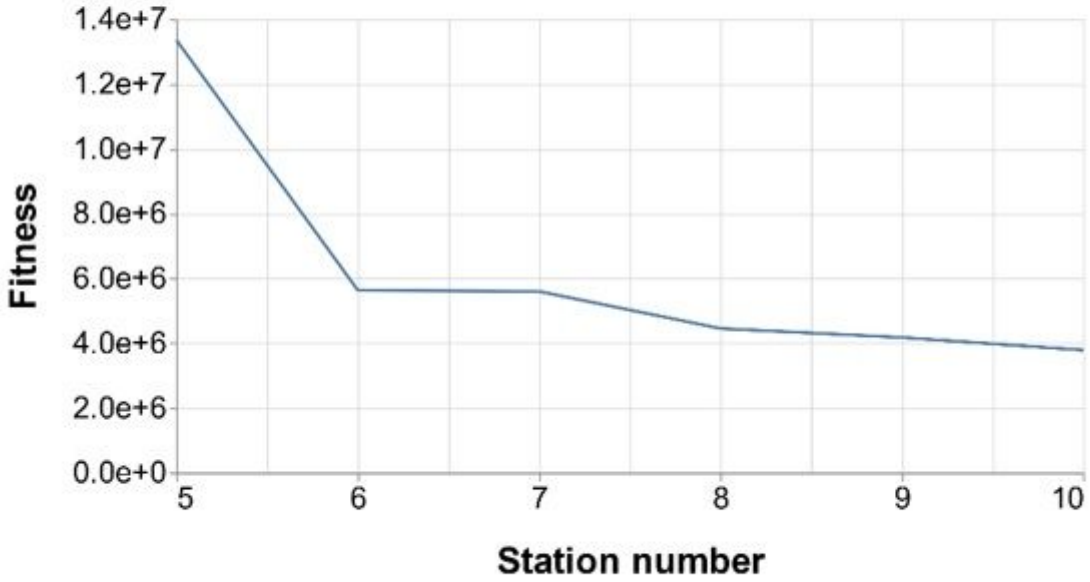


Figure 3

Visualization of event flow density, relief station location and supply volume in each station. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.



(a) Sensitivity analysis of total supply volume in Odaiba, Aug 3



(b) Sensitivity analysis of station number in Odaiba, Aug 3

Figure 4

Sensitivity analysis: Fitness functions in relation to station numbers and supply volumes