

Location Optimization of Urban Drone Delivery Cabinets Considering Three-Level Gradient Noise Penalty

Wenzhi Wang, Manzhen Duan*

School of Emergency Management and Safety Engineering, North China University of Science and Technology, Tangshan 063210, China

Abstract

This study puts forward a two-stage site selection method for courier cabinets integrated with a three-tier gradient noise penalty, aiming to resolve weak noise-sensitivity constraints and the unbalanced spatial layout of distribution facilities in suburban “drone delivery + self-pickup” systems. In the first stage, researchers use the K-means clustering method to merge massive delivery demand points into weighted demand nodes, and candidate facility locations are generated by gridding the study area. In the second stage, we construct a 0-1 integer programming model whose objective is to maximize weighted coverage after noise penalty adjustment. The established model adopts a three-tier gradient noise penalty function with threshold boundaries of 200 m and 500 m. It also enforces a minimum spacing requirement of 1800 m between any two distribution sites. A greedy algorithm equipped with an exchange operator considering noise penalties is adopted to solve this model. For empirical verification, we select 81,955 valid SF Express orders collected from Fengnan District, Tangshan City as the research case. Raw demand data are simplified into 200 weighted demand nodes, and 177 noise-sensitive points are added into the case analysis. The results indicate that the three schemes attain weighted coverage rates of 75.71%, 80.97%, and 83.55%, respectively. The marginal benefit brought by newly added distribution sites shows a declining trend. Besides, none of the three schemes arrange facilities inside the 200 m high-penalty zone; approximately 77% of selected sites fall within the 500 m non-affected region, and the overlap rate of service coverage circles stays below 15%. In general, the service-oriented scheme with nine distribution sites obtains the best overall performance in terms of coverage, construction cost and noise environmental suitability.

Keywords

Drone Delivery; Parcel Locker Site Selection; Noise-Sensitive Constraints; Three-Level Gradient Penalty; Two-Stage Heuristic Algorithm; Weighted Maximum Coverage.

1. Introduction

With the fast expansion of online shopping and changes in people's shopping behaviors, express parcels in China keep growing year by year, the old way of sending packages straight to customers' doors meets many troubles; labor expenses keep going up, and the delivery system can hardly handle orders during busy seasons. Serving as the final section of logistics work, last-mile parcel delivery takes up a large part of total delivery spending, so many firms try to work on this part to lower spending and raise working efficiency. Drone parcel delivery does not get stuck by road traffic jams, and it can adjust working plans easily, this method has been more and more used to finish parcel tasks in city areas.

However, drone rotor noise has become increasingly prominent and is a key factor limiting their large-scale application. Research by Zhou et al.[1] indicates that even small civilian drones can increase the regional background sound pressure level by at least 20 decibels. Through urban soundscape perception experiments, Torija et al.[2] further found that in areas with low traffic background noise, hovering drones can cause participants' noise annoyance levels to rise significantly from 2.3 to 6.8; the disturbance caused by drones in sensitive areas that are normally quiet-such as hospitals, schools, and nursing homes-is particularly pronounced. In urban planning practice, Zhou Yanan[3] applied acoustic environmental zoning and control techniques to differentiate and manage priority protection zones, key control zones, and general control zones, providing a reference framework for selecting noise-friendly sites for drone facilities.

In research on siting end-point facilities, traditional approaches to locating parcel lockers often use the set-coverage model or the maximum-coverage model to minimize the number of facilities, maximize coverage, or minimize delivery distances. Xiang Yanping et al.[4], using the Houjiatang District of Changsha as a case study, employed the set-coverage model to optimize the location and spatial distribution of logistics and express parcel lockers from the dual perspectives of delivery efficiency and customer satisfaction; Tian Shijia et al.[5] identified high-demand areas through demand assessment and combined clustering algorithms with heuristic search to achieve site selection with maximum coverage. Yang et al.[6] constructed a two-level planning model that divides communities into saturated and undersaturated zones to optimize the layout of new smart parcel lockers. Ghaderi et al.[7] treated parcel lockers as handover nodes for crowdsourced delivery, constructing a model for locker site selection and delivery task allocation and designing a two-stage solution algorithm. When it comes to solution algorithms, Zhou Jingwen[8] adjusted the K-means clustering method, adding demand weight values in its iteration steps, to find better positions for logistics distribution centers. Most of this existing work centered on regular manual package delivery cases, and they ignored noise created when drones take off and land; for this reason, the site selection models and limiting rules from these papers cannot be directly used when planning where to place drone parcel lockers.

For research about choosing locations of drone delivery equipment, Dukkanci and other authors[9] pointed out in their review paper that researchers have discussed many location problems in this area, including drone bases, charging points and material replenishment stations, yet there is still not enough work studying how takeoff-and-landing facility layouts interact with limits from the urban environment. Aurambout and co-workers[10] used high-precision population and land use data, to check possible positions and market reach for hive-style drone delivery stations inside European cities [11] established a planning model for the allocation of takeoff and landing points for last-mile delivery drones and solved it using a genetic algorithm; Zhang Honghai et al.[12] further considered the hierarchical classification of takeoff and landing points, no-fly zones, and the alignment of supply and demand to construct a multi-level layout for urban logistics drone takeoff and landing points aimed at minimizing economic costs and maximizing customer satisfaction. Xue Qingwen et al.[13] proposed a method for selecting sites for hierarchical drone takeoff and landing fields and established a multi-objective maximum-coverage site selection model aimed at maximizing demand coverage while minimizing construction costs and overlap rates. Qi et al.[14] combined K-means clustering with the Gray Wolf Optimization algorithm to construct a site selection model for urban last-mile joint distribution centers that integrates air-ground coordination; Liu et al.[15] established a hierarchical site selection and allocation model, using a two-layer hybrid heuristic algorithm combining genetic algorithms and simulated annealing to determine the optimal number and locations of sites at each level; Lu Lingling and Hu Zhihua[16] designed a two-stage algorithm combining K-means clustering with an improved simulated annealing method to solve the site selection problem for emergency delivery on islands; Liu Kanglin et al.[17] employed a distribution-robust optimization method to characterize the impact of demand uncertainty on the site selection and allocation of UAV delivery facilities; Wan Lili et al.[18] constructed a fuzzy maximum coverage model for emergency scenarios, using triangular fuzzy numbers to characterize the uncertainty of coverage radii.

Furthermore, to address the semi-enclosed nature of residential communities, Li Zhuolun et al.[19] used the FCM algorithm to identify ideal takeoff and landing points that balance flight safety and noise impact. While these studies offer valuable insights for UAV facility site selection, they lack a refined, gradient-based characterization of noise impact severity and pay little attention to spatial balance among delivery locker locations.

In summary, while existing research has yielded substantial results in the site selection of last-mile facilities and the layout of UAV takeoff and landing points, the following shortcomings remain: First, most studies employ fixed buffer zones to address noise constraints, lacking a graded and refined characterization of noise impact; second, site selection objectives often focus on a single dimension—such as coverage or cost—while insufficiently considering the spatial balance of locker locations; third, research on noise-friendly site selection for the emerging last-mile model of "drone delivery combined with locker pickup" remains insufficient. Accordingly, this paper uses Fengnan District, Tangshan City, as the study area and develops a site selection model for urban last-mile drone parcel lockers that accounts for noise-sensitive constraints. The study first uses the K-means clustering algorithm to reduce the dimensionality of parcel demand points, generating a weighted set of demand points; it then establishes a two-stage site selection model that integrates a three-level gradient noise penalty, minimum inter-locker spacing constraints, and weighted demand coverage. A noise-penalty greedy algorithm combined with a permutation operator for local search is designed to solve this model; finally, the model's effectiveness is validated through a comparative analysis of multiple scenarios, providing a decision-making reference for the scientific layout of urban last-mile drone parcel lockers.

2. Model Formulation

2.1 Problem Description

This paper investigates how to optimize site selection for urban last-mile drone delivery lockers under noise constraints. Given parcel demand points, noise-sensitive facilities, and candidate locker locations within the study area, the decision-maker must select several candidate locations to deploy drone parcel lockers. While satisfying service radius, minimum spacing, and noise-sensitivity constraints, the decision-maker must determine a reasonable number of locker locations and their spatial layout to maximize the weighted total covered parcel demand.

Specifically, the study focuses on a defined region. After applying the K-means clustering algorithm to reduce the dimensionality of the original parcel demand points within the region, N weighted demand points are generated $D_{\min} = \{d_1, d_2, \dots, d_N\}$. Each demand point has latitude and longitude coordinates and a shipment volume weight w_i . The total weight is 1, representing the proportion of parcels within a given cluster, and it characterizes differences in demand intensity across areas. For noise-sensitive facilities, this paper collected and organized five types of noise-sensitive public service facilities within the study area: hospitals, nursing homes, universities, middle schools, and elementary schools. After cleaning the spatial data and standardizing the coordinates, these were aggregated into M noise-sensitive points $S = \{s_1, s_2, \dots, s_M\}$, as spatial reference points subject to noise constraints. The Set of candidate cabinet points $C = \{c_1, c_2, \dots, c_K\}$. We generated this set by rasterizing the study area and using regular grid nodes as candidate cabinet locations.

Each parcel locker has a fixed service radius R ; a demand point is considered to be covered only if the spherical distance from that point to the locker location does not exceed R . To avoid wasting service resources due to excessive concentration of locker locations, the spherical distance between any two locker locations must not be less than the minimum spacing threshold $D_{\min}$. Regarding noise constraints, a three-level gradient noise penalty mechanism is employed: different penalty coefficients are assigned based on the distance from a candidate cell location to the nearest noise-sensitive point; the closer the distance, the heavier the penalty. This guides the algorithm, through the objective function, to prioritize candidate locations that are farther from sensitive areas. This paper

aims to determine a set of cabinet locations $L \subseteq C$ that maximize the weighted demand coverage rate after adjusting for noise penalties, while analyzing differences in site selection among three scheme types—cost-oriented, service-oriented, and coverage-oriented-by varying cabinet size parameters. The problem can be formally expressed as in Equation (1):

$$Max = \frac{\sum_{i=1}^N w_i \cdot y_i \cdot \max_{j \in L} \{\eta_j \cdot z_{ij}\}}{\sum_{i=1}^N w_i} \quad (1)$$

2.2 Definition of Symbol

Table 1. Definition of Model Symbols

Category	Symbol	Definition
Set	D	Weighted Set of demand points $D = \{d_1, d_2, \dots, d_N\}$
	S	Set of noise-sensitive points, $S = \{s_1, s_2, \dots, s_M\}$
	C	Set of candidate cabinet locations $C = \{c_1, c_2, \dots, c_K\}$
	L	Set of final selected cabinet locations, $L \subseteq C$
	I	Set of original parcel pickup points, used for the first stage of clustering
Parameters	N	Number of Weighted Demand Points After Clustering
	M	Number of Noise-Sensitive Points
	K	Number of Candidate Locker Locations
	p	Number of Deployed Locker Locations
	R	Parcel Locker Service Radius (m)
	$D_{\min}$	Minimum Distance Threshold Between Lockers (m)
	d_1, d_2	Noise Penalty Threshold (m) $d_1 < d_2$
	α, β, γ	Level-3 noise penalty coefficient $\alpha < \beta < \gamma \leq 1$
	w_i	Weight of parcel volume at demand point i
	K_c	Number of clusters in the first-stage K-means clustering
	x_j	0-1 variable: 1 if candidate point j is selected as a locker point; 0 otherwise
	y_i	0-1 variable: 1 if demand point i is covered; 0 otherwise
Intermediate variable	z_{ij}	0-1 variable: 1 if demand point i is covered by locker point j; 0 otherwise
	η_j	Noise penalty coefficient for locker point j $\eta_j \in \{\alpha, \beta, \lambda\}$
	δ_{ij}	Spherical distance from demand point i to locker point j (m)
	τ_j	Distance from locker point j to the nearest noise-sensitive point (m)

Model BuThis paper develops a two-stage site selection model. In the first stage, the K-means clustering algorithm is used to perform spatial dimensionality reduction on the massive dataset of raw parcel demand points, generating a weighted set of demand points; in the second stage, a maximum coverage site selection model is established that integrates a three-level gradient noise penalty, inter-locker spacing constraints, and a weighted coverage objective.

2.2.1 Dimension Reduction Using K-Means Clustering

Let D denote the Set of initial parcel demand points. $I = \{q_1, q_2, \dots, q_{N_0}\}$ Each point has latitude and longitude coordinates. The goal of K-means clustering is to partition N_0 the original data points into K_c clusters such that the sum of the squared distances from each point to its cluster center is minimized.

The objective function is:

$$\min J = \sum_{k=1}^{K_c} \sum_{q_i \in G_k} \|q_i - \mu_k\|^2 \quad (2)$$

Here, G_k is the k th cluster, μ_k is the center coordinates of the k th cluster, and $\|q_i - \mu_k\|$ is the spherical distance between the two points.

After clustering is complete, each cluster center serves as a weighted demand point, with a weight equal to the number of original demand points in that cluster:

$$w_k = |G_k|, k = 1, 2, \dots, K_c \quad (3)$$

$$d_k = \mu_k, k = 1, 2, \dots, K_c \quad (4)$$

This generates a weighted set of demand points $D = \{d_1, d_2, \dots, d_N\}$ $N = K_c$.

2.2.2 Noise-Penalized Maximum Coverage Site Selection Model

The objective function is:

$$Max = \frac{\sum_{i=1}^N w_i \cdot y_i \cdot \eta_i^*}{\sum_{i=1}^N w_i} \quad (5)$$

Here, η_i^* is the noise penalty coefficient corresponding to the cabinet location covering demand point i , defined as:

$$\eta_i^* = \max_{j \in L} \{\eta_j \cdot z_{ij}\} \quad (6)$$

The numerator is the weighted total of covered demand points, adjusted for noise penalties, and the denominator is the weighted total of all demand points.

The economic interpretation of the objective function is as follows: Under cabinet placement scheme L, cover as many high-weight demand points as possible while prioritizing cabinet locations that are far from noise-sensitive areas.

(1) Coverage Constraint: A demand point i is covered if and only if there exists at least one cabinet point j such that the spherical distance between the two points does not exceed the service radius R , and that cabinet point is selected:

$$\begin{aligned} y_i &\leq \sum_{j=1}^K z_{ij}, \forall i \in D \\ z_{ij} &\leq x_j, \forall i \in D, \forall j \in C \\ z_{ij} &\leq 1(\delta_{ij} \leq R), \forall i \in D, \forall j \in C \end{aligned} \quad (7)$$

In Equation (7), $z_{ij} \leq 1(\delta_{ij} \leq R)$ is an indicator function that equals 1 when the condition in parentheses is true and 0 otherwise.

Unique allocation constraint: Each demand point may be served by at most one cabinet location:

$$\sum_{j=1}^K z_{ij} \leq 1, \forall i \in D \quad (8)$$

Cabinet-to-cabinet spacing constraint: The spherical distance between any two selected cabinet points must not be less than the minimum spacing threshold $D_{\min}$:

$$\delta_{jj'} \geq D_{\min} \cdot (x_j + x_{j'} - 1), \forall j, j' \in C, j \neq j' \quad (9)$$

The meaning of the constraint in Equation (9) is that when both candidate points are selected $(x_j + x_{j'} - 1)$, their distance must be equal to $\delta_{jj'} \geq D_{\min}$; when at least one is not selected, the constraint is automatically relaxed.

(4) Cabinet-point quantity constraint: The total number of selected cabinet points equals the plan's set value p :

$$\sum_{j=1}^K x_j = p \quad (10)$$

(5) Noise penalty constraint: Based on the distance τ_j from the cabinet location to the nearest noise-sensitive point, the penalty coefficient is determined according to a three-level gradient:

$$\eta_j = \begin{cases} \alpha, \tau_j < d_1 \\ \beta, d_1 \leq \tau_j < d_2 \\ \gamma, \tau_j \geq d_2 \end{cases} \quad (11)$$

In Equation (11), $\tau_j = \min_{s \in S} \delta(c_j, s)$ is the minimum spherical distance from cabinet point j to all noise-sensitive points.

This paper uses $d_1 = 200\text{m}, d_2 = 500\text{m}, \alpha = 0.3, \beta = 0.7, \gamma = 1.0$. Their physical meanings are as follows: When the distance between a cabinet location and a noise-sensitive point is less than 200 m, the noise from UAV takeoffs and landings severely disturbs the sensitive area; the penalty coefficient is set to 0.3, meaning that the coverage contribution of that cabinet location is counted as only 30% in the objective function; When the distance is between 200 and 500 m, the noise impact is moderate, and the penalty coefficient is set to 0.7; when the distance exceeds 500 m, the noise impact is negligible, and the penalty coefficient is set to 1.0, with no reduction applied. A schematic of the three-level gradient noise penalty function is shown in Figure 1.

The three-level gradient noise penalty function proposed in this paper is a continuous cost mechanism: candidate points are not directly eliminated simply because they are close to sensitive regions; instead, their contribution to the objective function is reduced by lowering their penalty coefficients. The algorithm automatically balances these factors during point selection: if a candidate point has exceptionally strong coverage capability, its overall score may still be higher than that of a candidate point farther from the sensitive area but with weaker coverage capability, even if the former is closer to the sensitive area and has a lower penalty coefficient; conversely, if coverage capabilities are similar, the algorithm prioritizes the candidate point with the higher penalty coefficient. This mechanism dynamically balances coverage and environmental impact. A schematic of the three-level gradient noise penalty function is shown in Figure 1.

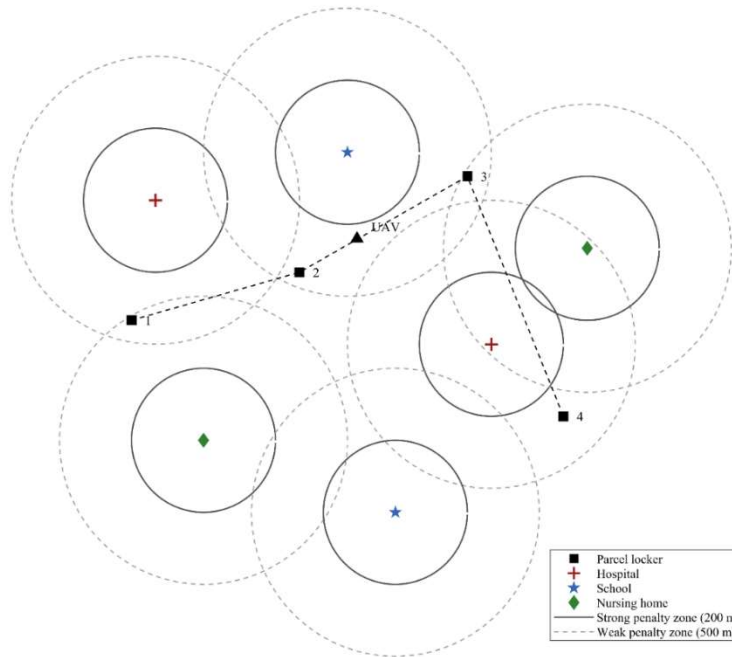


Figure 1. Schematic Diagram of a Three-Level Gradient Noise Penalty

2.2.3 Calculation of Spherical Distance

The spherical distance between any two points $a = (lon_a, lat_a)$ and $b = (lon_b, lat_b)$ is calculated using the Haversine formula:

$$\delta(a, b) = 2R_e \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{lat_b - lat_a}{2}\right) + \cos(lat_a)\cos(lat_b)\sin^2\left(\frac{lon_b - lon_a}{2}\right)}\right) \quad (12)$$

In Equation (12), $R_e = 6371000m$ represents the average radius of the Earth, and both latitude and longitude are expressed in radians.

3. Algorithm Design

The site selection model described above is a 0-1 integer programming model with nonlinear constraints, which belongs to the class of NP-hard problems. When the number of candidate cabinet locations is large, solving the problem exactly becomes difficult; therefore, this paper proposes a two-stage heuristic algorithm. The main thought behind this algorithm works like this: in the first stage, we apply K-means clustering to cut down spatial dimensions for the large original group of demand points, which creates a new group of weighted demand points, and we build the candidate cabinet position group by dividing the research area into grid cells; in the second stage, we adopt the greedy algorithm with noise penalty to get a basic site arrangement plan, and we carry out repeated local search improvements through a swap operator to adjust this plan. Figure 2 shows the whole structure of this algorithm.

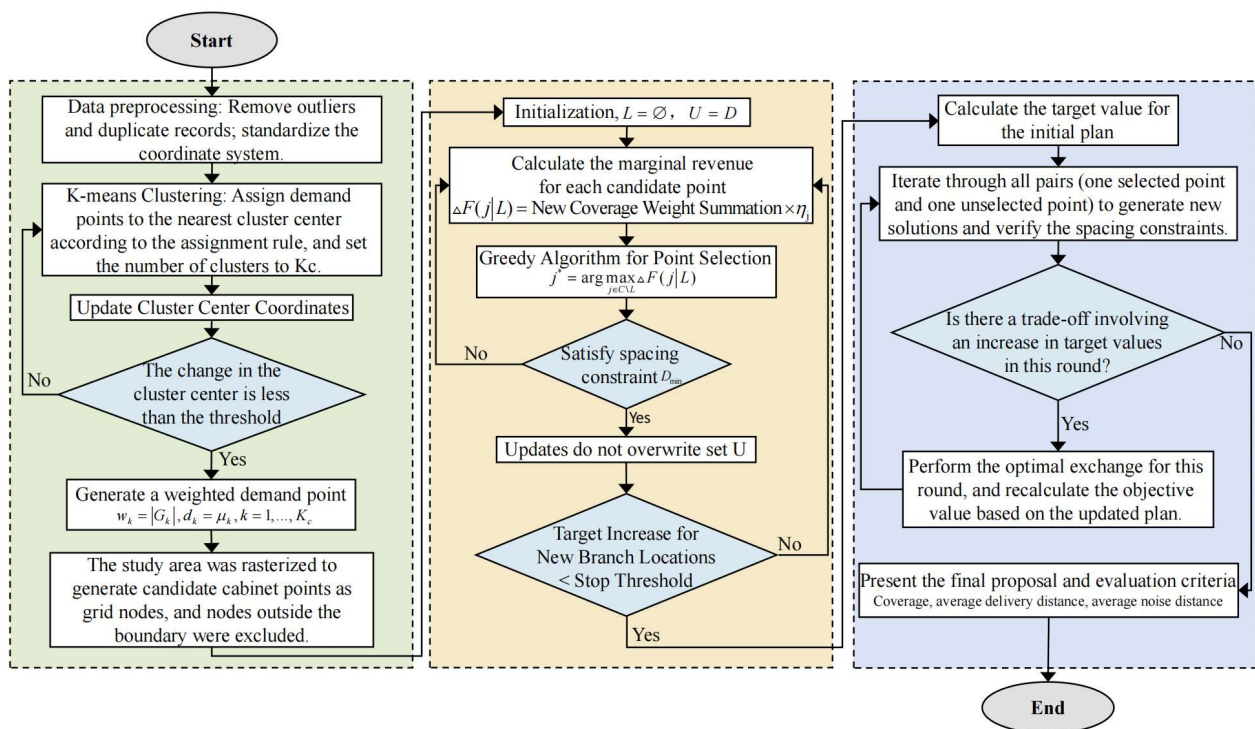


Figure 2. Flowchart of the Two-Stage Heuristic Algorithm

3.1 Phase 1: Dimension Reduction via Clustering of Demand Points and Generation of Candidate Points

Phase 1 is meant to shrink the huge raw parcel demand point dataset into a smaller set of weighted demand points that are easier to handle, and it also creates candidate locker positions spread out in regular patterns, which supplies basic data for the site selection optimization work of Phase 2.

Step 1: Data Preprocessing

We tidy up the original parcel demand points, removing abnormal records and repeated latitude and longitude entries, and we unify the coordinate format; we also combine noise-sensitive points and delete repeated entries, so both groups of spatial data can use the same coordinate system.

Step 2: Generate Weighted Demand Points Using K-Means Clustering

Set the number of clusters K^c and use the K-means algorithm to cluster the preprocessed demand points spatially. Each demand point q_i is assigned to the nearest cluster center according to the following assignment rule:

$$q_i \in G_k^* k^* = \arg \min_{k \in \{1, \dots, K_c\}} \delta(q_i, \mu_k) \quad (13)$$

Here, G_k^* represents the k-th cluster, μ_k represents the center coordinates of the k-th cluster, and $\delta(q_i, \mu_k)$ represents the spherical distance function.

The rule for updating the cluster center is the arithmetic mean of the coordinates of all points within the cluster:

$$\mu_k^{(t+1)} = \frac{1}{|G_k^{(t)}|} \sum_{q_i \in G_k^{(t)}} q_i \quad (14)$$

Iterate until the change in the cluster center falls below the threshold ε or the maximum number of iterations $T_{\max}$ is reached. Once clustering is complete, each cluster center is treated as a weighted demand point, with a weight equal to the number of original demand points in that cluster:

$$w_k = |G_k|, d_k = \mu_k, k = 1, \dots, K_c \quad (15)$$

Step 3: Generate candidate cabinet locations by gridding the study area.

Using the study area's administrative boundaries as the scope, generate a regular grid at fixed intervals, with the grid nodes serving as candidate cabinet locations. After excluding grid nodes that fall outside the administrative boundaries, the Set of candidate cabinet locations $C = \{c_1, c_2, \dots, c_K\}$ is obtained. For each candidate cell location, the distance τ_j to the nearest noise-sensitive point is calculated in advance, and its penalty coefficient η_j is determined based on a three-level gradient noise penalty function:

$$\tau_j = \min_{s \in S} \delta(c_j, s), \eta_j = \begin{cases} \alpha, \tau_j < d_1 \\ \beta, d_1 \leq \tau_j < d_2 \\ \gamma, \tau_j \geq d_2 \end{cases} \quad (16)$$

Precomputing $\tau_j \eta_j$ avoids redundant calculations in the second-stage iteration, improving the algorithm's efficiency.

3.2 Phase 2: Noise-Penalized Greedy Algorithm

In the second stage, using the weighted demand points and candidate cabinet locations generated in the first stage as input, an initial site selection plan is constructed using a noise-penalized greedy algorithm. In each round, the greedy algorithm selects the candidate cabinet location with the highest marginal benefit from the Set of candidate locations. It adds it to the Set of selected locations until it reaches the specified number of cabinet locations, p.

Step 1: Calculate the marginal benefit of a candidate point.

For any unselected candidate point c_j , its marginal benefit is the additional weighted total demand points covered after adding c_j to the current set of selected points L , adjusted for the noise penalty. Specifically, it is calculated as follows:

$$\Delta F(j|L) = \sum_{i \in U(L)} w_i \cdot 1(\delta_{ij} \leq R) \cdot \eta_j \quad (17)$$

Here, $U(L)$ represents the Set of demand points not covered by the selected set L , and $1(\delta_{ij} \leq R)$ is an indicator function. The physical interpretation is: the sum of the weights of demand points that the candidate point c_j can cover-and that other cabinet points have not yet covered-multiplied by that candidate point's noise penalty coefficient. The lower the penalty coefficient η_j , the lower the marginal benefit, thereby discouraging the algorithm from selecting points near sensitive areas.

Step 2: Greedy Point Selection

In each round, select the candidate point with the highest marginal gain and add it to the Set of selected points:

$$j^* = \arg \max_{j \in C \setminus L} \Delta F(j|L) \quad (18)$$

Step 3: Distance Constraint Check

After selecting a candidate point j^* , you must verify whether the distances between it and all cabinet points in the selected Set satisfy the minimum distance constraint:

$$\delta(j^*, j) \geq D_{\min}, \forall j \in L \quad (19)$$

If the condition is not met, the candidate point is discarded, and the candidate point with the next-highest marginal benefit is selected for testing. After satisfying the spacing constraint, j^* it is added to the selected set L , and the Set of uncovered demand points $U(L)$ is updated.

Step 4: Termination of Iteration

Repeat Steps 1 through 3 until the number of selected cabinet locations reaches the set value p , and output the initial site selection plan $L^{(0)}$.

3.3 Local Search Using the Exchange Operator

Because the greedy algorithm's initial solution is prone to getting stuck in a local optimum, this paper introduces a local search using a swap operator to improve it iteratively. The swap operator repeatedly attempts to replace a selected cabinet with an unselected candidate; if the objective function improves after the replacement, the swap is accepted; otherwise, the original solution is retained.

Step 1: Initial Solution and Objective Value Calculation

Using the initial solution $L^{(0)}$ output by the greedy algorithm as the starting point for the local search, calculate its objective function value:

$$F(L) = \frac{\sum_{i=1}^N w_i \cdot y_i \cdot \eta_i^*}{\sum_{i=1}^N w_i} \quad (20)$$

Here, y_i and η_i^* represent, respectively, the coverage status of demand point i under Scheme L and the corresponding noise penalty coefficient.

Step 2: Swap Operation

For each pair of (j_{out}, j_{in}) , where $j_{out} \in L$ is a selected cabinet point and $j_{in} \in C \setminus L$ is an unselected candidate point, perform a swap operation to generate a new solution:

$$L' = (L \setminus \{j_{out}\}) \cup \{j_{in}\} \quad (21)$$

After the exchange, verify whether the new plan meets the cabinet spacing constraints. If it does not, the exchange is not feasible and should be skipped.

Step 3: Acceptance Criteria

Calculate the objective function value for the new plan L' . If $F(L')$ is better than $F(L)$, it does, then accept the swap and let $L' \rightarrow L$; otherwise, reject the swap and retain the original plan.

Step 4: Termination of Iteration

Iterate through all feasible swap pairs. If a swap is accepted in this round, return to Step 2 to continue the next search round; if no swap is accepted, the solution has reached a local optimum, the algorithm terminates, and the final site selection solution L^* is output.

To improve search efficiency, this paper adopts an optimal improvement strategy: in each round, all feasible swap pairs are traversed, and the swap that maximizes the objective function is selected for execution, rather than accepting the first improvement encountered. This strategy reduces the number of iteration rounds and accelerates convergence.

3.4 Overall Process and Iterations

The overall execution process of the algorithm is as follows:

- 1) Execute Phase 1: Perform K-means clustering on the original demand points to reduce dimensionality, generating a weighted set of demand points; simultaneously, rasterize the study area to generate a set of candidate cabinet locations and pre-calculate the noise penalty coefficients for each candidate location.
- 2) Execute the second stage-the noise-penalty greedy algorithm-selecting sites in successive rounds according to the marginal benefit maximization criterion to generate the initial site selection plan $L^{(0)}$.
- 3) We implement local search on this initial solution through the exchange operator. The algorithm repeatedly attempts to substitute selected sites with unused candidate locations until the solution converges to a local optimum.
- 4) Output the final site selection plan L^* along with its objective function value, coverage rate, average delivery distance, and other evaluation metrics.

4. Case Studies

4.1 Overview of the Study Area and Data Processing

4.1.1 Overview of the Study Area

This study chooses Fengnan District of Tangshan City as the research area for empirical analysis. Fengnan District sits on the Bohai Sea coast in eastern Hebei Province and belongs to the southern

urban zone of Tangshan's city center. It has thickly distributed residents and clustered housing estates, which creates strong demand for last-mile parcel delivery. This region contains many noise-sensitive buildings such as hospitals, nursing facilities, as well as primary and secondary schools; these features make it a representative case to examine the parcel locker location model with noise-related restrictions. The study area spans approximately 117.80° ~ 118.45° E and 39.17° ~ 39.66° N. Spatial processing and visualization were performed in QGIS 3.3. Model solving and algorithm implementation were performed in MATLAB R2024b.

4.1.2 Data Sources and Preprocessing

The express delivery demand data used in this study was derived from actual delivery orders processed by SF Express's Fengnan District in April 2026. The raw dataset contained 86,301 records, each with the latitude and longitude coordinates of the express packages. We obtained noise-sensitive location data through the Baidu Maps Open Platform's POI interface, covering five categories of public service facilities: hospitals, nursing homes, universities, middle schools, and elementary schools. Data on the administrative boundaries of the study area were sourced from the National Geographic Information Resource Catalog Service System.

Data preprocessing primarily included the following steps: ① Coordinate system unification—converting all spatial data to the GCJ-02 coordinate system to ensure that parcel points, sensitive points, and administrative boundaries share a consistent coordinate reference frame; ② Outlier removal—deleting records with latitude and longitude coordinates outside the administrative boundaries of Fengnan District, with coordinates equal to zero, or with obvious coordinate drift; ③ Duplicate record removal: Multiple orders at the same coordinate location were retained for subsequent weighting calculations, but completely duplicate fields were removed. After cleaning, 81,955 valid parcel records and 177 noise-affected sensitive points were obtained.

4.1.3 Data Sources and Preprocessing

The original dataset contains over 80,000 parcel points. If these were directly used as input for the site selection model, the resulting distance matrix would be too large, leading to inefficient solution calculation. This paper uses the K-means algorithm to cluster the demand points and reduce dimensionality. The number of clusters K_c is selected using the elbow rule: the sum of squared errors within clusters (SSE) is used as the metric to evaluate clustering performance, and its calculation formula is as follows:

$$SSE = \sum_{k=1}^{K_c} \sum_{q_i \in G_k} \|q_i - \mu_k\|^2 \quad (22)$$

Here, μ_k represents the k-th cluster, and μ_k is the center of that cluster. As K_c increases, SSE continues to decrease, but the rate of decrease slows significantly after a certain value, forming an "elbow" inflection point. In this paper, multiple experiments were conducted with K_c ranging from 50 to 400. When $K_c = 200$, the decline in SSE leveled off, and increasing the number of clusters further yielded only limited improvements in clustering accuracy; therefore, the optimal number of clusters was 200.

After clustering, we take each cluster centroid as one weighted w_k demand point. Its weight value corresponds to the count of original parcel entries inside the corresponding cluster, which reflects the local demand intensity. This procedure finally yields 200 weighted demand points; the dataset size falls to 0.24% of the initial more-than-80,000 records, and the general spatial pattern of demand distribution remains well preserved. Figure 4 presents the geographic positions and weight distribution of demand points after clustering. Points with large weights mainly gather within the

built-up zone in central Fengnan District, matching the real distribution of population and commercial activities.

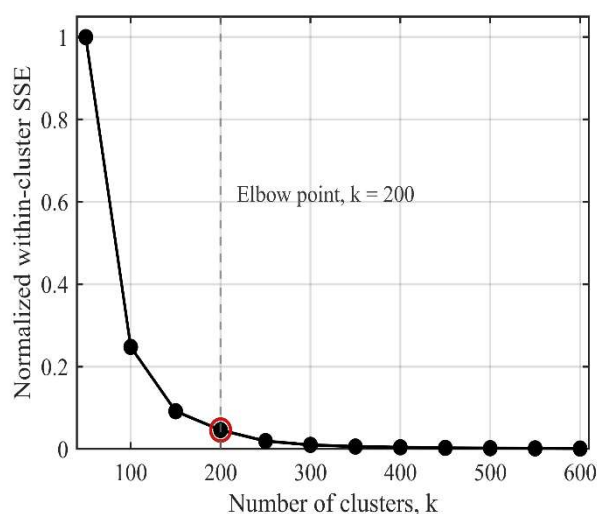


Figure 3. Determining the Optimal Number of Clusters Using the Elbow Method

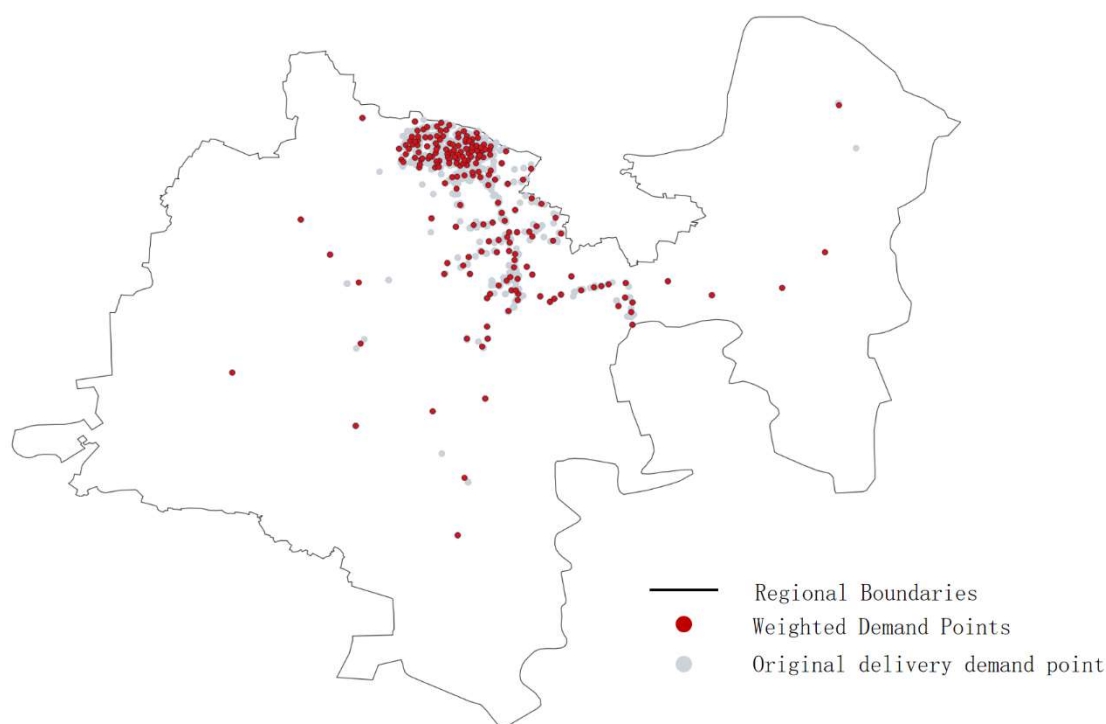


Figure 4. Distribution of weighted demand points after clustering

4.1.4 Generation of Noise-Sensitive Points and Candidate Cabinet Locations

The POI data covering five categories, namely hospitals, nursing homes, universities, high schools and primary schools, were spatially integrated and deduplicated. A total of 177 noise-sensitive points were obtained to act as spatial references for the three-level gradient noise penalty.

Candidate cabinet sites are produced through rasterization of the study area. Regular grids with fixed spatial resolution are constructed within the administrative scope of Fengnan District, and grid nodes are set as candidate positions for parcel lockers. Nodes beyond administrative boundaries are eliminated to form the final candidate set. The grid resolution balanced site-selection accuracy with computational efficiency: too high a resolution would create redundant candidate points and increase solution time. At the same time, too low a resolution might omit high-quality locations. Following

preliminary experiments, this study adopted a grid spacing that matches the service radius. For each candidate cabinet location, we pre-calculated the spherical distance $d_{\min}(j)$ to distance to the nearest noise-sensitive point. We determined the noise penalty coefficient η_j to avoid redundant calculations during the algorithm's iterative process.

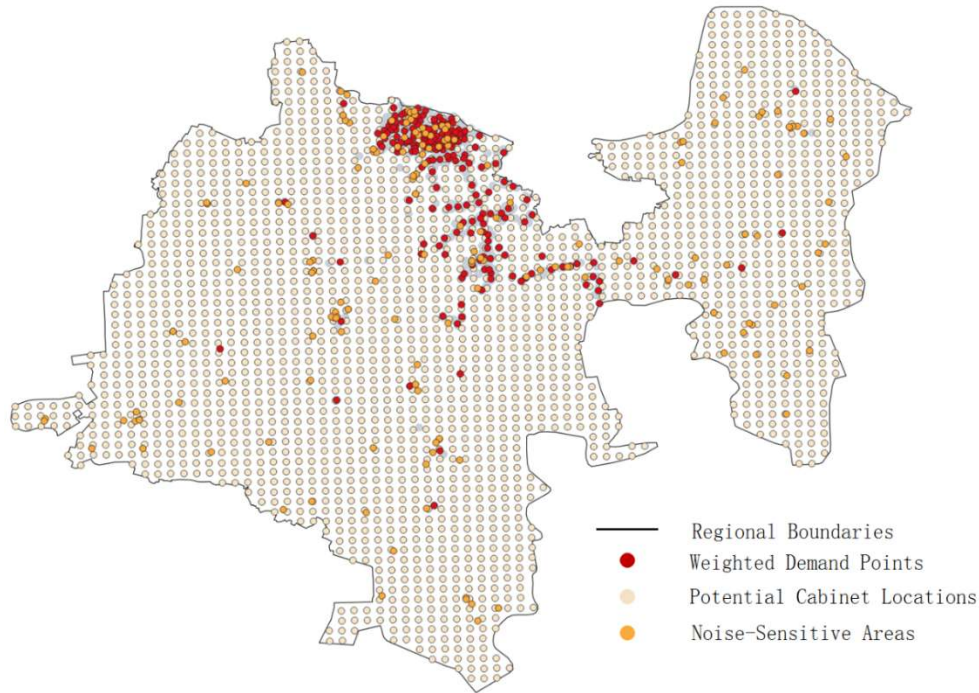


Figure 5. Rasterization of the Study Area and Spatial Distribution of Site Selection Factors

4.2 Parameter Settings

Parameter values adopted in the model are summarized in Table 1. The service radius R of parcel lockers is defined as 1,000 m; this value matches a walking distance of roughly 10 minutes for residents to collect packages, and it conforms to the common service coverage of urban self-pickup facilities for last-mile logistics. The minimum distance $D_{\min}$ between lockers is set at 1,800 m, about 1.8 times the service radius, to minimize overlap between adjacent locker service areas and avoid redundant resource allocation. The noise penalty thresholds d_1 and d_2 are set at 200m and 500m, corresponding to strong and weak noise impact ranges; the three-level penalty coefficients α , β , and γ are set at 0.3, 0.7, and 1.0, respectively. The number of clusters k is set to 200.

Table 2. Model Parameter Values

Parameters	Meaning	Value
R	Parcel Locker Service Radius / m	1000
$D_{\min}$	Minimum Distance Between Lockers / m	1800
d_1	Strong Penalty Distance Threshold / m	200
d_2	Weak Penalty Distance Threshold / m	500
α	Strong Penalty Coefficient	0.3
β	Weak Penalty Coefficient	0.7
γ	No Penalty Coefficient	1.0
K_c	Number of Clusters in K-means Clustering	200

4.3 Analysis of Site Selection Results

4.3.1 Sensitivity Analysis of the Number of Branches

The number of cabinet locations directly determines construction costs and service coverage levels, making it a core variable in site selection decisions. This paper does not preset a single number of cabinet locations; instead, it sets the number of cabinet locations, p , to take sequential values within the range $1 \sim 30$. For each value of p , a two-stage heuristic algorithm is run to obtain the corresponding weighted demand coverage rate, and a sensitivity curve showing how the coverage rate varies with the number of cabinet locations is plotted, as shown in Figure 6.

The results indicate that when the number of cabinet locations is small ($p < 6$), the coverage rate rises rapidly as the number of cabinet locations increases, with each additional cabinet location yielding a significant gain in coverage; when p falls within the range $6 \sim 13$, the growth rate of the coverage rate slows but continues to rise steadily; when p exceeds 13, the curve flattens, and the marginal contribution of new cabinet locations to coverage decreases significantly. Adding more cabinet locations primarily results in overlapping service areas rather than covering new demand areas. Considering the marginal trade-off between increased coverage and construction costs, this paper selects three representative cabinet point scales, $p = 6$, $p = 9$, and $p = 13$, corresponding to cost-oriented, service-oriented, and coverage-oriented approaches, respectively, for subsequent comparative analysis. Specifically, $p = 6$ is at the end of the rapid growth phase, with the primary objective of controlling construction costs; $p = 13$ is near the curve's saturation inflection point, aiming to maximize coverage; and $p = 9$ lies between the two, balancing coverage levels with resource investment.

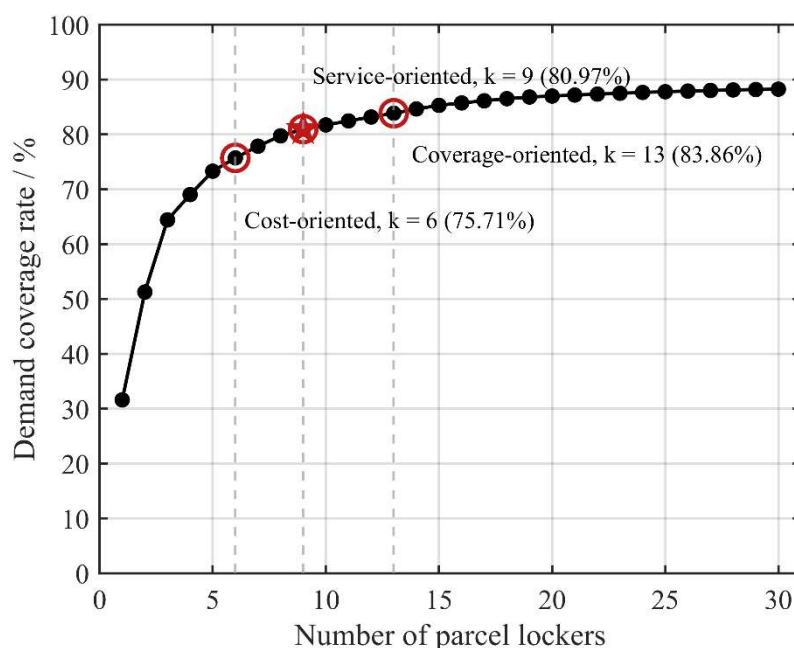


Figure 6. Sensitivity Analysis of the Number of Package Lockers vs. Coverage Rate

4.3.2 Comparison of Site Selection Results for the Three Options

Table 3 compares the evaluation metrics for the three plan types. The evaluation metrics include weighted demand coverage, the average distance from a demand point to the nearest cabinet location, the average distance from a cabinet location to the nearest noise-sensitive point, and the number of sensitive points within a 500-meter radius of a cabinet location.

Table 3. Evaluation Criteria for the Three Types of Site Selection Plans

Solution	Number of Lockers	Weighted Coverage Rate (%)	Average Delivery Distance (m)	Average Noise Distance (m)	Service Area Overlap Rate (%)
Cost-Oriented	6	75.71	690.4	577.4	6.7
Service-Oriented	9	80.97	682.8	716.3	5.6
Coverage-Oriented	13	83.55	657.0	790.6	12.8

Fengnan District's express delivery demand presents an obvious spatial pattern of single-core agglomeration. The weighted demand points after clustering gather heavily within the central built-up zone. By comparison, the northern and southern portions of the study region are dominated by extensive towns and farmland, where delivery demand barely exists. If the map were plotted based on the full administrative boundary, the core zone for site selection would appear squeezed, and large blank regions would occupy the map. This hinders clear observation of the spatial connections between locker sites and demand points. For this reason, the research narrows the mapping scope to the central built-up area where demand and locker candidates concentrate. Figure 7 displays site selection outcomes from the three groups of schemes. Within this figure, red dots stand for weighted demand points, and the size of each dot reflects the parcel volume weight of its corresponding cluster; blue circles mark parcel lockers; pale green triangles denote noise-sensitive points; and dashed circles illustrate service radii.

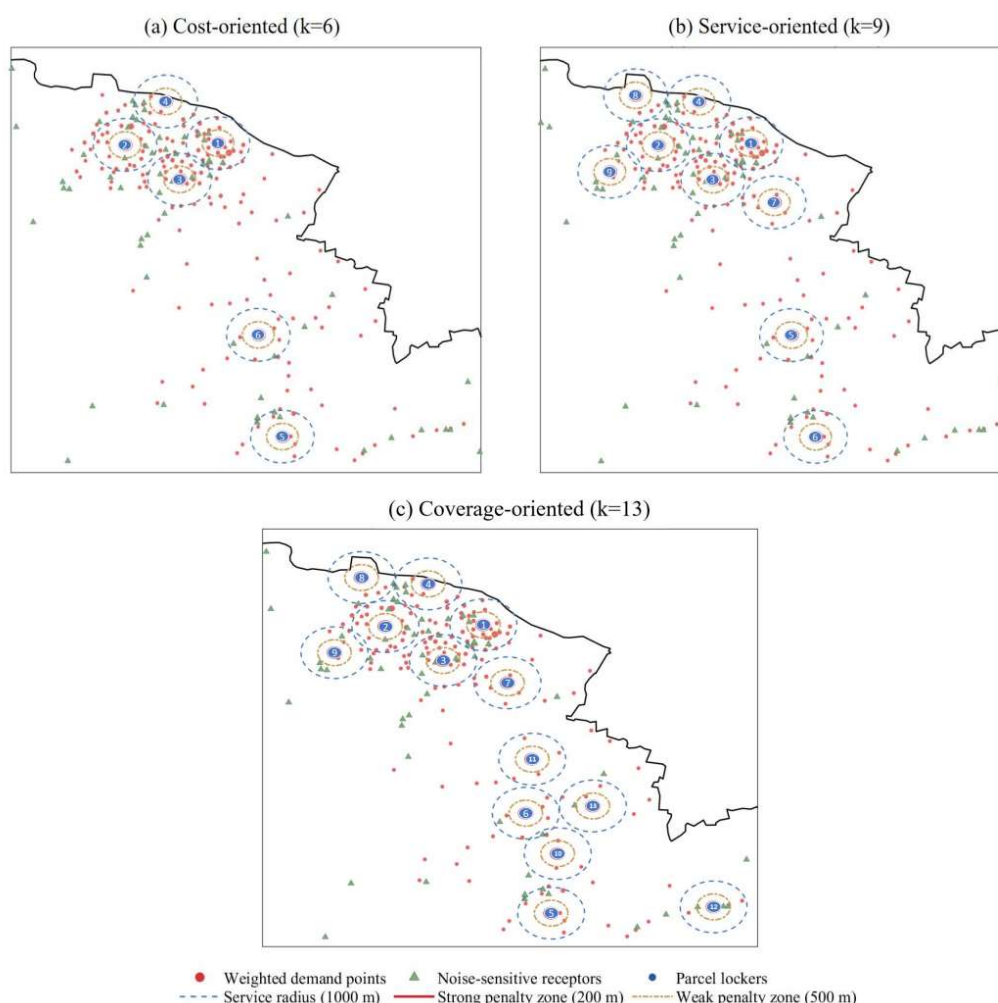


Figure 7. Spatial layout results for three representative schemes

(1) Cost-Oriented Approach. This strategy employs the minimal number of locker sites, which brings the lowest expenses for construction and operation, and reaches a weighted coverage rate of 75.71%. Locker facilities are mainly arranged inside several core clusters carrying the largest demand weights, and each site takes charge of a high amount of service demand. Due to the small total number of lockers, certain peripheral demand points lie beyond the service radius. Besides, to cover high-demand regions, a small number of lockers sit relatively close to noise-sensitive zones. This scheme fits the initial phase of delivery pilot projects with restricted construction budgets.

(2) Service-Oriented Scheme. This option strikes a balance between cost and coverage, attaining a weighted coverage rate of 80.97%. It rises by 5.26 percentage points compared with the cost-oriented scheme, while merely three extra lockers are added. These newly-added lockers efficiently fill coverage gaps on the outer edge of the core region. Meanwhile, more available candidate sites allow the algorithm to select positions farther away from noise-sensitive regions, so this scheme achieves better overall noise adaptability than the cost-oriented one. It keeps coverage, construction expenditure and environmental influence in balance; for this reason, this study treats it as the recommended solution.

(3) Coverage-Oriented Scenario. This case contains the largest quantity of locker sites and obtains the highest weighted coverage rate of the three schemes at 83.55%. Its service zones can basically cover major demand regions within the study boundary. Nevertheless, relative to the service-oriented scheme, adding four more lockers only lifts the coverage rate by 2.58 percentage points, which demonstrates the law of diminishing marginal returns. Moreover, partial service zones of these lockers overlap with each other. This scenario applies to regions that require high service coverage and have adequate construction capital.

In terms of spatial layout, lockers from all three schemes are distributed along corridors with dense demand. None of these selected sites fall into the 200-meter strict penalty zone around noise-sensitive regions. This outcome proves that the three-tier noise penalty mechanism provides valid guidance for site selection results. As the total number of lockers grows, newly selected sites gradually extend toward peripheral areas with lower demand density and more desirable environmental conditions.

5. Conclusion

This paper addresses the urban last-mile delivery scenario of "drone delivery combined with parcel locker pickup." To address issues such as insufficient characterization of noise-sensitive constraints in parcel locker site selection, lack of spatial balance among locker locations, and difficulties in solving for large-scale demand points, we construct a weighted maximum coverage site selection model that incorporates a three-level gradient noise penalty. We design a two-stage heuristic algorithm comprising "K-means clustering for dimensionality reduction-noise-penalized greedy site selection-exchange operator local search" and conduct an empirical analysis based on actual parcel data and noise-sensitive point data from SF Express's Fengnan District in April 2026. The main conclusions are as follows:

(1) The two-stage clustering-greedy site selection model put forward in this research can well handle location optimization problems involving large-scale demand datasets. At the first stage, K-means clustering reduces more than 80,000 original parcel records into 200 weighted demand points. This step greatly cuts down the problem scale while keeping the spatial distribution pattern of demand intact; at the second stage, site searching is completed through a noise-penalized greedy algorithm coupled with local search built on the permutation operator. The solving procedure avoids random initialization, so the produced results remain stable and repeatable. This method offers a viable technical route for locating last-mile service facilities when facing massive order datasets.

(2) A three-tier gradient noise penalty function is adopted to substitute conventional rigid buffer zones, which creates a good trade-off between service coverage and noise environmental suitability. Instead of directly ruling out candidate sites close to sensitive regions, this mechanism assigns differentiated cost values according to distance. It guides the optimization model to actively stay away from

sensitive establishments including hospitals, schools and nursing homes. Calculation outcomes reveal that roughly 77% of selected lockers fall inside the 500-meter no-impact zone. The mechanism lifts noise adaptability substantially without sacrificing overall coverage performance.

(3) Introduction of the minimum separation constraint between lockers improves the spatial layout balance effectively. The overlap ratio of service coverage between neighboring lockers drops from about 45% (in the unconstrained case) to below 15%. This prevents lockers from gathering too tightly and stops unnecessary waste of service resources. After comprehensive evaluation across indicators such as coverage rate, average delivery distance, average noise separation distance and service overlap rate, the deployment of nine parcel lockers is suggested for Fengnan District. This configuration reaches the best compromise among construction expenditure, service quality and environmental influence.

This study uses point features to approximate noise-sensitive facilities. The model does not account for real building boundaries or how drone flight trajectories affect noise spread. Subsequent research may introduce building outline datasets and three-dimensional noise decay models to realize refined noise constraints at the flight-path level. Besides, time-varying demand data can be incorporated to construct dynamic location models and strengthen the layout's capacity to cope with changing demand volumes.

Acknowledgments

This work was supported by the Tangshan Science and Technology Plan Project (Grant No. 25130209B) and the Research Project of China Federation of Logistics & Purchasing (Grant No. 2026CSLKT3-443).

The authors sincerely thank Associate Professor Manzhen Duan for her precious guidance, thoughtful comments and consistent support during the whole research process. Her professional knowledge and encouragement played an important role in finishing this manuscript.

References

- [1] Zhou, Z., & Brandão, M. (2023). Noise and environmental justice in drone fleet delivery paths: a simulation-based audit and algorithm for fairer impact distribution. In *Proceedings of IEEE International Conference on Robotics and Automation*. IEEE.
- [2] Torija, A. J., Li, Z., & Self, R. H. (2020). Effects of a hovering uncrewed aerial vehicle on urban soundscapes perception. *Transportation Research Part D: Transport and Environment*, 78, 102195.
- [3] Zhou, Y. N., Fan, D. P., Zhang, Y. K., et al. (2025). Exploration of technologies and practices for acoustic zoning and control. *Journal of Environmental Engineering Technology*, 15(1), 38–47.
- [4] Xiang, Y. P., & Lei, A. H. (2022). Site selection and spatial optimization of logistics and express pickup lockers from the perspectives of delivery efficiency and customer satisfaction: A case study of Houjiatang North, Changsha. *Economic Geography*, 42(8), 87–94.
- [5] Tian, S. J., Bao, D. W., & Cheng, H. (2023). Evaluation of demand for airport bus stops in urban areas and methods for their site selection. *Transportation Research*, 9(2), 22–32.
- [6] Yang, X., Wang, C., He, X., et al. (2023). Location optimization for community smart parcel lockers based on bilevel programming. *Journal of Advanced Transportation*, 2023, 1998188.
- [7] Ghaderi, H., Zhang, L., Tsai, P. W., et al. (2022). Crowdsourced last-mile delivery with parcel lockers. *International Journal of Production Economics*, 251, 108549.
- [8] Zhou, J. W. (2026). A study on the optimization of distribution center site selection based on an improved K-Means algorithm. *China Logistics and Purchasing*, (7), 35–36.
- [9] Dukkanci, O., Campbell, J. F., & Kara, B. Y. (2024). Facility location decisions for drone delivery: a literature review. *European Journal of Operational Research*, 316(2), 397–418.
- [10] Aurambout, J. P., Gkoumas, K., & Ciuffo, B. (2019). Last mile delivery by drones: an estimation of viable market potential and access to citizens across European cities. *European Transport Research Review*, 11, 30.

- [11] Qian, X. Y., Zhang, H. H., Zhang, F., et al. (2021). A study on the site selection and allocation of takeoff and landing points for last-mile delivery drones. *Journal of Wuhan University of Technology (Transportation Science and Engineering Edition)*, 45(4), 682–687, 693.
- [12] Zhang, H. H., Feng, D. K., Zhang, X. W., et al. (2022). A study on the planning of takeoff and landing site layouts for urban logistics drones. *Transportation Systems Engineering and Information*, 22(3), 207–214.
- [13] Xue, Q. W., Jiang, Y., Feng, Y. L., et al. (2026). A study on the site selection of tiered takeoff and landing sites for UAVs in low-altitude logistics. *Journal of South China University of Technology (Natural Science Edition)*, 54(6), 29–41.
- [14] Qi, W., Li, A., & Zhang, H. (2024). Location selection methods for urban terminal co-distribution centers with air–land collaboration. *Applied Sciences*, 14(13), 5814.
- [15] Liu, Z., Nie, L., Xu, G., et al. (2023). Multi-objective design of UAS air route network based on a hierarchical location–allocation model. *Sustainability*, 15(23), 16521.
- [16] Lu, L. L., & Hu, Z. H. (2022). Site selection and route optimization for drone delivery relay stations on islands. *Journal of Dalian University of Technology*, 62(3), 299–308.
- [17] Liu, K. L., Zhang, Z. Y., Jiang, J. W., et al. (2026). The site selection and allocation problem for UAV delivery facilities based on distribution-robust optimization. *Journal of Jilin University (Engineering Edition)*, 56(2), 464–472.
- [18] Wan, L. L., Xu, S. M., Huang, J. H., et al. (2026). A study on the site selection of emergency delivery centers for UAVs based on a fuzzy maximum coverage model. *Transportation Systems Engineering and Information*, 26(3), 166–175.
- [19] Li, Z. L., Lu, J., Wang, X. R., et al. (2026). A method for the coordinated site selection of urban logistics drone takeoff and landing points and truck stopping points. *Journal of Transportation Engineering*, 26(3), 89–105.