

Article

Spatial Allocation of Veterinary Healthcare Facilities in Beijing Using Random Forest and Priority-Constrained Covering Optimisation

Boting Xi ¹, Shaohua Wang ^{1,2,*} , Siyu Zhao ³, Wenda Wang ³, Liang Long ⁴, Chang Liu ^{1,2}, Jingyi Zhou ¹ and Haojian Liang ¹ 

¹ State Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China; xibt21@lzu.edu.cn (B.X.); liuchang248@mails.ucas.ac.cn (C.L.); zhoujingyi@aircas.ac.cn (J.Z.); lianghj@aircas.ac.cn (H.L.)

² Aerospace Information Research Institute, University of Chinese Academy of Sciences, Beijing 100049, China

³ Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China; 11240888@stu.lzjtu.edu.cn (S.Z.); wangwd@lzjtu.edu.cn (W.W.)

⁴ Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; longl@reis.ac.cn

* Correspondence: wangshaohua@aircas.ac.cn

Abstract

Veterinary hospitals are increasingly important neighborhood services, but citywide evidence on their strategic allocation remains limited. Existing GIS-based suitability studies identify favorable areas, but do not determine which limited set of planning nodes to prioritise when existing supply is clustered and underserved areas require explicit protection. This study develops a transparent allocation framework for Beijing by integrating Random Forest (RF)-based observed clinic-location probability, E2SFCA-derived service-gap indicators, and a priority-constrained covering optimisation model. The RF output is interpreted as a proxy for location attractiveness, rather than a direct estimate of demand for veterinary healthcare. The optimisation then protects predefined priority areas and maximises weighted coverage under alternative facility budget scenarios. Additional analyses include a threshold-derived priority set, class-imbalance metrics, catchment-radius sensitivity, key-area-count sensitivity, systematic candidate/key-score coefficient perturbation, district-grouped spatial validation, raw versus calibrated RF allocation sensitivity, and pet-service endogeneity checks. Under the main 5 km catchment and $P = 40$ scenario, the model achieved complete priority-area coverage and weighted coverage of 0.656. Coverage increased further at higher budgets, so $P = 40$ is not regarded as a practical optimum, but rather as a diminishing-return scenario. At $P = 20$ and $P = 40$, priority-area coverage increased by 5.0 and 2.5 percentage points relative to MCLP, respectively, while aggregate weighted coverage remained comparable. The framework supports strategic planning priority analysis, not direct commercial site selection.

Keywords: veterinary hospitals; location-allocation; priority-constrained covering model; random forest; spatial accessibility; Beijing



Academic Editor: Wolfgang Kainz

Received: 6 July 2026

Revised: 2 September 2026

Accepted: 4 September 2026

Published: 11 September 2026

Copyright: © 2026 by the authors.

Published by MDPI on behalf of the International Society for

Photogrammetry and Remote

Sensing. Licensee MDPI, Basel,

Switzerland. This article is an open

access article distributed under the

terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/)

license.

1. Introduction

Urban pet healthcare is increasingly becoming an important component of community services, yet access to veterinary services remains spatially uneven. Existing studies

have identified veterinary service deserts and documented significant geographical disparities in access to veterinary clinics, demonstrating that accessibility is contingent upon the combined distribution of providers, households, transportation networks, and socioeconomic conditions [1,2]. These relationships are further complicated in megacities like Beijing, which have a dense urban core, a rapidly urbanizing periphery, and significant disparities between residential and commercial areas. Thus, the planning challenge is not solely determined by the presence of veterinary hospitals in a district, but rather by the potential for additional capacity to enhance access without duplicating the existing concentrated supply.

Demand for veterinary hospitals cannot be represented by population size alone. The use of companion animal healthcare is shaped by owner information, household circumstances, clinical decision-making, and the role of pets in everyday family life [3–6]. Vaccination, preventive care, chronic disease management, and age-related treatment also generate recurring service needs throughout an animal's life [7,8]. Accordingly, pet healthcare is a neighborhood-scale accessibility issue with both welfare and market dimensions: potential users require reasonable proximity to services, while private clinics also depend on viable residential demand and a supportive environment of complementary services.

From a planning perspective, existing approaches address distinct components of this challenge, but fail to comprehensively integrate spatial diagnosis, facility selection, and priority protection. GIS-based suitability assessment and multicriteria decision-making enable the identification of spatially favorable locations by synthesizing demographic, accessibility, land-use, environmental, infrastructure, and feasibility data [9–11]. However, a ranked suitability surface alone cannot determine which limited set of facilities to select under a fixed budget, nor explicitly accommodate overlapping service coverage areas. Conventional covering models provide a more direct allocation framework. The location set covering problem seeks to minimize the number of facilities required to achieve a pre-specified coverage threshold. In contrast, the maximal covering location problem (MCLP) maximises covered demand under a fixed facility budget [12,13]. Nevertheless, an unconstrained MCLP may favor dense, high-weight locations and leave smaller or geographically dispersed underserved areas insufficiently protected. This inherent trade-off between efficiency and equity is especially salient in Beijing, where high-demand central areas coexist with peripheral locations where veterinary services are less accessible. To address this problem, this study develops an integrated RF-SWCLM (Random Forest Stage-Weighted Covering Location Model) framework that links observed clinic-location patterns, service-gap diagnosis, candidate-site screening, and priority-constrained facility allocation. Random Forest (RF) estimates the probability of observed veterinary hospital locations using heterogeneous urban spatial covariates. Importantly, this probability is interpreted as a proxy for location attractiveness rather than for direct veterinary healthcare demand. RF was selected not because it is assumed to be universally superior to alternative algorithms such as XGBoost or LightGBM, but because it provides a transparent and reproducible means of modeling nonlinear relationships while supporting permutation-based feature interpretation, spatial validation, class-imbalance assessment, and probability-calibration diagnostics. The resulting spatial signal is subsequently integrated with population, residential intensity, service accessibility, and feasibility information for allocation modeling.

The optimisation component adopts a priority-constrained covering structure in which predefined underserved priority areas are protected before the remaining facility budget is used to improve aggregate weighted coverage. Rather than treating this structure as a fundamentally new optimisation algorithm, this study reformulates its demand representation, candidate-generation procedure, priority-area definition, service-gap indicators, and planning interpretation in the context of veterinary healthcare. The result-

ing model therefore represents a planning objective that differs from conventional MCLP rather than claiming universal superiority in aggregate coverage. The main contributions of this study are summarized as follows:

- Observed-location-based planning-proxy construction. An RF model is employed to characterize the spatial distribution of existing veterinary hospitals using covariates related to population, residential, pet-service, commercial, accessibility, green-space, and feasibility. The resulting RF output is explicitly interpreted as a planning proxy for the observed attractiveness of clinic locations rather than a direct measure of pet ownership or healthcare utilization.
- Integration of potential demand and service gaps. RF-derived spatial signals, population intensity, and residential intensity, E2SFCA (Enhanced Two-Step Floating Catchment Area)-based accessibility, and existing coverage gaps are integrated to identify locations where relatively strong demand potential coincides with inadequate service provision.
- Metropolitan candidate and priority screening. Feasible candidate planning nodes and predefined key priority areas are screened across Beijing's heterogeneous 16-district metropolitan environment by jointly considering demand-related signals, service gaps, buildability, and accessibility.
- Priority-constrained allocation and benchmark evaluation. The inherited priority-constrained covering framework is adapted to veterinary healthcare allocation and compared with MCLP under alternative facility budgets and modeling assumptions. This comparison enables explicit evaluation between aggregate weighted coverage and guaranteed coverage of priority areas.

These contributions constitute a domain-specific methodological adaptation and empirical extension of an existing covering-location framework rather than the development of a fundamentally new covering-location algorithm.

The empirical analysis yields three decision-relevant findings. Existing veterinary hospitals exhibit a pronounced concentration in central and inner-suburban service clusters. The RF model identifies pet-related service density, population density, and residential density as important predictors of the observed probability of clinic location. Still, these associations should not be interpreted as evidence of causal demand generation. Under the main 5 km catchment scenario, the $P = 40$ allocation achieves complete priority-area coverage while capturing most of the attainable weighted coverage within the tested candidate set. Accordingly, $P = 40$ is interpreted as an elbow-type planning scenario rather than as a mathematically optimal facility number or a directly implementable commercial target.

2. Related Work

Spatial planning for urban service facilities has been approached primarily through suitability assessment, accessibility analysis, machine-learning-based spatial prediction, and location-allocation optimisation. These methodological streams address complementary aspects of facility planning, but each has limitations when applied independently to the allocation of veterinary healthcare at the metropolitan scale.

GIS-based suitability assessment and multicriteria decision-making are widely used to identify spatially favorable locations for healthcare and other urban facilities. These approaches integrate demographic, accessibility, land-use, environmental, infrastructure, and feasibility indicators to generate interpretable suitability surfaces [9–11]. Their principal advantage is transparency in spatial screening; however, suitability rankings alone do not determine how many facilities should be selected under a fixed resource constraint, explicitly account for overlapping service areas among candidate sites, or ensure that underserved locations will be included in the final allocation. Suitability analysis is therefore

more appropriate for identifying potentially favorable areas than for solving the complete facility-allocation problem.

Machine learning and GeoAI approaches provide complementary capabilities for representing complex spatial relationships from heterogeneous urban data. Points of Interest (POIs) provide detailed information on the spatial distribution of urban activities and service facilities, whereas OpenStreetMap provides road-network, building, and land-use information, although its completeness may vary by location and feature type [14,15]. Random Forest models can capture nonlinear relationships between observed facility locations and heterogeneous spatial covariates [16]. In the present study, RF is used to characterize the spatial signature of existing veterinary hospital locations rather than to directly estimate pet ownership, healthcare demand, or clinical utilization. Alternative boosting algorithms such as XGBoost and LightGBM may also achieve strong predictive performance, but relative advantages depend on the data structure, tuning procedure, and evaluation objective. Because the machine learning output is subsequently incorporated into the allocation framework as a planning proxy, interpretability, spatial validation, probability diagnostics, and reproducibility are important alongside predictive discrimination. RF was therefore selected as an interpretable and sufficiently robust predictive component rather than on the assumption that it universally outperforms alternative machine learning models.

Spatial accessibility research provides another important methodological foundation for facility planning. Healthcare accessibility reflects interactions among population distribution, provider capacity, transportation barriers, facility hierarchies, competition, and travel impedance [17–20]. Studies involving traffic-sensitive healthcare access, vaccination services, and elderly care further show that catchment definitions and distance-decay assumptions can substantially influence the identification of underserved populations [21–23]. Floating-catchment approaches are particularly useful because they simultaneously account for spatial supply–demand relationships and travel impedance. Nevertheless, accessibility indicators are primarily diagnostic: they identify areas where service provision is relatively strong or weak; they do not directly determine which combination of new facilities should be selected under a constrained facility budget.

A broader body of GIS-based planning and GeoAI research also demonstrates the value of integrating heterogeneous spatial evidence into urban decision-making. GIS-AHP and related multicriteria techniques provide structured approaches for combining suitability and feasibility indicators [24,25]. At the same time, GeoAI and geospatial information research emphasize the increasing role of data-driven spatial analysis in urban planning [26,27]. Related applications have combined multicriteria analysis, open urban data, land-use information, and machine learning to characterize suitable facility locations and complex urban spatial structures [28–30]. These studies demonstrate the analytical value of combining heterogeneous urban datasets, but translating such evidence into discrete facility-allocation decisions remains a distinct optimisation problem.

Research on accessibility inequality and proximity-based urban planning further emphasizes that the provision of facilities should not be evaluated solely by administrative counts. Studies of pedestrian accessibility, inclusive 15 min city planning, and access to public services and infrastructure demonstrate that spatial proximity and distributional inequality can remain substantial even where aggregate service provision appears sufficient [31–33]. This insight is directly relevant to veterinary healthcare planning because a large number of existing clinics does not necessarily imply adequate coverage of underserved neighborhoods.

Location-allocation models provide a more direct mechanism of translating spatial evidence into facility decisions. Classical covering models identify facility configurations

according to explicit coverage objectives. MCLP is particularly useful when resources are limited and overlapping facility coverage must be considered simultaneously [12,13]. However, because its objective emphasizes aggregate covered demand, an unconstrained MCLP may favor high-weight areas and leave smaller or spatially dispersed underserved areas uncovered. Thus, maximising overall coverage does not necessarily ensure protection of strategically important priority areas.

The closest methodological precedent for addressing this limitation is the SWCLM developed for waste-bin allocation in the Mount Yun Scenic Area [34]. That study considered a bounded tourist environment, represented heterogeneous waste-generation demand, guaranteed coverage of key locations, and subsequently improved general coverage. Its staged priority logic provides a useful mechanism for balancing aggregate efficiency with explicit protection of selected areas. However, direct application of this structure to metropolitan veterinary healthcare would be inappropriate because waste bins are standardized public environmental facilities. In contrast, veterinary hospitals operate within an open urban service market characterized by heterogeneous residential demand, complementary pet-service ecosystems, existing competition, land constraints, and uncertain service capacity.

Additional research on healthcare accessibility, green-space accessibility, road-network vulnerability, and proximity-based urban service planning further demonstrates that identification of underserved areas depends strongly on catchment definitions, travel conditions, urban morphology, and service-network completeness [35–44]. These studies reinforce the need to conceptualize veterinary hospital allocation as a spatial accessibility and service-planning problem rather than as a simple count of existing facilities. At the same time, these accessibility-oriented approaches primarily diagnose spatial inequalities and do not by themselves determine the optimal combination of new facilities under a fixed planning budget.

The methodological gap does not lie in the absence of suitability mapping, accessibility analysis, machine learning, or the individual coverage of optimisation. Rather, relatively limited attention has been given to integrating these components within a transparent metropolitan veterinary healthcare allocation framework in which service gaps are explicitly diagnosed, candidate locations are screened for planning feasibility, and underserved priority areas are protected under a fixed facility budget. The present study addresses this gap by combining RF-based observed clinic-location probability, E2SFCA-derived service-gap indicators, feasibility-based candidate screening, and priority-constrained covering optimisation. In doing so, it adapts the inherited SWCLM structure for an open metropolitan veterinary-service context while retaining the MCLP as a benchmark for evaluating the trade-off between aggregate coverage efficiency and priority-area assurance. The major methodological approaches and their general limitations are summarized in Table 1.

Table 1. Comparison of major approaches for spatial facility planning and allocation.

Approach	Main Capability	Limitations
GIS suitability/GIS-MCDM [9–11,24,25]	Integrates demographic, accessibility, environmental, land-use, and feasibility criteria	Sensitive to criterion selection and weighting, and does not directly produce optimised allocation decisions
MCLP [12,13]	Maximises covered demand under a fixed facility budget while accounting for overlapping coverage	Depends on predefined coverage thresholds and may overlook equity and heterogeneous service capacity

Table 1. Cont.

Approach	Main Capability	Limitations
Machine learning/GeoAI [16,26–30]	Captures nonlinear relationships between facility locations and heterogeneous urban covariates	Depends on training-data quality and spatial representativeness, and does not directly yield allocation decisions
Spatial accessibility/E2SFCA [17–23,31–33,35–44]	Evaluates supply–demand relationships, travel impedance, and service inequality	Sensitive to catchment and distance-decay assumptions, and is primarily evaluative rather than prescriptive
SWCLM [34]	Introduces priority-area assurance before general coverage improvement	Depends on priority-area definitions and stage-specific constraints, and may simplify heterogeneous service conditions

3. Study Area and Data

3.1. Study Area

The study area is Beijing Municipality, China, which comprises 16 district-level administrative units. Beijing provides an appropriate case for research on veterinary hospital allocation because of its highly heterogeneous urban structure, including a dense urban core, mature inner-suburban residential and commercial districts, new-town development zones, and extensive peripheral ecological and mountainous areas. These pronounced spatial contrasts result in substantial variation in population density, road accessibility, existing veterinary hospital provision, developable land availability, and potential service gaps. A 1 km grid was adopted as the basic analytical unit to balance spatial resolution, citywide coverage, computational efficiency, and compatibility with population and POI-derived indicators. This resolution is consistent with the spatial granularity of the WorldPop gridded population dataset and reduces excessive sparsity in POI-derived indicators at finer spatial resolutions. All distance, area, and buffer analyses were performed using the WGS 84/UTM Zone 50N coordinate reference system (EPSG:32650), which is suitable for Beijing's longitude range. Final maps were exported using standardized boundary symbology, legends, north arrows, and non-overlapping scale bars. To assess whether the identified priority pattern was sensitive to spatial resolution, the priority scores were additionally aggregated to a 2 km grid and compared with the 1 km results at the district level. Table 2 presents the grid-resolution sensitivity analysis for the 1 km and 2 km grids. The Spearman correlation of district-level priority counts was 0.851, indicating that the overall spatial pattern remains largely stable under the coarser-grid scenario (Figure 1).

Table 2. Grid-scale sensitivity check.

Check	Value	Interpretation
1 km vs. 2 km priority-count Spearman rho	0.851	District-level priority pattern remains similar after 2 km aggregation
Haidian	10 vs. 11	1 km key areas vs. 2 km priority cells
Dongcheng	8 vs. 6	1 km key areas vs. 2 km priority cells
Xicheng	7 vs. 4	1 km key areas vs. 2 km priority cells
Fengtai	6 vs. 7	1 km key areas vs. 2 km priority cells
Shijingshan	4 vs. 4	1 km key areas vs. 2 km priority cells

Note: The 2 km check aggregates the priority-score components to coarser grid cells and compares district-level priority counts with the 1 km result.

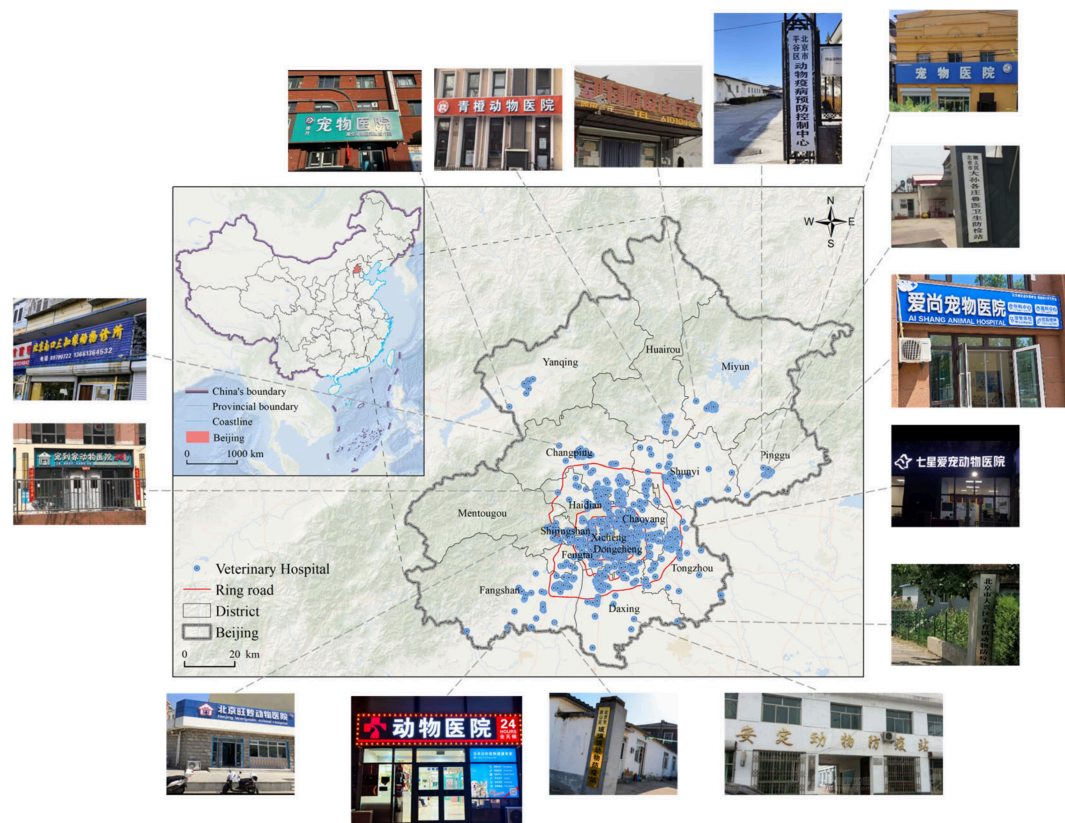


Figure 1. Study area and distribution of veterinary hospitals in Beijing. (Images of veterinary hospitals from AMap). Note: The Chinese text in the photographs is part of the original veterinary hospital signage. “宠物/动物医院” translates as “Veterinary Hospital.” The remaining Chinese characters indicate the specific hospital name.

3.2. Data Sources

This study integrates multisource geospatial datasets obtained from open-access repositories and web-based services. Data on veterinary hospitals, pet-related service facilities, residential communities, commercial centers, parks and green spaces, bus stops, and metro stations were collected from the AMap Web Service API. Population data were obtained from the WorldPop 2020 1 km gridded population dataset. Road networks, green spaces, water bodies, protected areas, and land-use information were derived from OSM (OpenStreetMap). Beijing municipal and district administrative boundaries were obtained from the DataV GeoAtlas boundary dataset. All spatial layers were harmonized to a common Beijing administrative boundary and projected into a consistent coordinate reference system to support distance, area, and overlay analyses (Table 3).

Table 3. Data sources and analytical roles.

Dataset	Source	Role in Analysis	Temporal Note
Administrative boundaries	DataV GeoAtlas	Municipal and district clipping, aggregation, and mapping	Cached and processed in June 2026
POI variables	AMap Web Service API	Veterinary hospitals, pet-related services, residential communities, commercial centers, parks, bus/metro stations	Collected and cleaned in June 2026

Table 3. Cont.

Dataset	Source	Role in Analysis	Temporal Note
Population	WorldPop 2020	Baseline grid-level population density and demand proxy	2020 gridded product; interpreted with 2026 POI/residential indicators
Roads and land-use constraints	OpenStreetMap/Geofabrik	Road density, major-road proximity, green-area ratio, and constrained land	Downloaded and processed in June 2026
Manual POI verification	Baidu Maps spot-check	Independent check of cleaned veterinary hospital POIs	30 records, 29 correct and 1 misclassification

Note: POI and OSM-derived layers were collected and processed in June 2026. Raw POI records are subject to AMap Web Service API terms.

The data integration strategy is consistent with established principles of spatial accessibility analysis and the use of open urban data. The original and enhanced 2SFCA models demonstrate the need to evaluate supply, demand, and travel impedance together, rather than relying only on administrative facility ratios [45–48]. WorldPop provides a transparent gridded demographic baseline [49], while studies of OSM data quality support the use of volunteered geographic information for road-network analysis when completeness limitations are explicitly acknowledged and assessed [50,51].

The population layer is a gridded dataset from 2020, whereas POI and OSM-derived variables were collected and processed in June 2026. The population variable is therefore interpreted as a baseline spatial distribution rather than a direct representation of the 2026 population distribution. To reduce the limitations of relying on an older population layer, the demand proxy also incorporates 2026 residential-community POIs and RF-derived location probabilities. This approach explicitly acknowledges the temporal limitation of the population data while using more recent urban-service indicators to update the demand context.

3.3. Spatial Covariates

The spatial covariates were selected to represent both observed location preferences and planning feasibility. Population density and residential density serve as demographic proxies for potential household-level demand. Pet-related service density captures the complementary service ecosystem, including pet retail stores, grooming, boarding, training, and veterinary-related stores. Commercial density approximates urban activity and the concentration of consumer services. Road density and transit accessibility represent general urban accessibility conditions. Green-space indicators reflect the availability of outdoor environments that are potentially relevant to pet-related activities. Service-gap indicators identify areas where relatively strong demand-related signals coincide with limited existing service accessibility (Figure 2, Table 4).

Figure 3 shows the dimension-level Spearman correlations among demand, accessibility, service ecology, coverage gap, and planning priority. The strongest positive correlation was observed between demand and planning priority ($\rho = 0.72$), whereas demand and accessibility showed the strongest negative correlation ($\rho = -0.71$). Accessibility was also negatively correlated with planning priority ($\rho = -0.65$), while service ecology showed a moderate positive correlation with coverage gap ($\rho = 0.53$). Overall, the absolute correlation coefficients ranged from 0.34 to 0.72, indicating moderate interrelationships without near-perfect correlations among the five dimensions. This pattern is consistent with the VIF (Variance Inflation Factor) results, with most variables exhibiting VIF values below 2.5, suggesting limited multicollinearity among the model inputs. These correlations are

interpreted as descriptive diagnostics of potential redundancy rather than inferential tests because the spatial grid observations cannot be assumed to be fully independent.

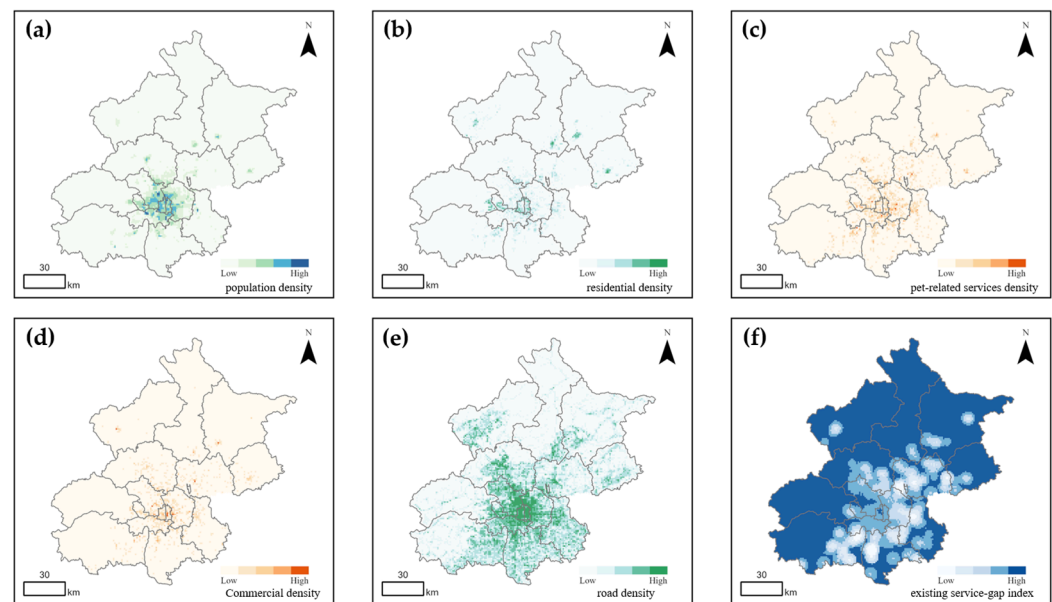


Figure 2. Input spatial covariates for allocation modeling. The panels show (a) population density, (b) residential density, (c) pet-related service density, (d) commercial density, (e) road density, and (f) the existing service-gap index.

Table 4. Indicator dimensions and model roles.

Dimension	Representative Variables	Model Role	Expected Interpretation
Demand	Population density, residential density, RF-derived planning weight	Demand-point weighting and candidate screening	Higher values indicate stronger potential veterinary healthcare demand
Service ecology	Pet-related service density, commercial density	RF features and planning-context indicators	Higher values indicate a stronger complementary service environment
Accessibility	Road density, distance to transit, distance to major roads	RF features and candidate eligibility	Better accessibility supports practical facility use
Coverage gap	E2SFCA service-gap index, existing hospital coverage gap	Key-area screening and scenario interpretation	Higher values indicate underserved demand potential
Planning priority	Buildable score, water/protected/constrained land ratio, priority-area score	Candidate screening and priority-zone interpretation	Higher values indicate a stronger planning priority after feasibility screening

Note: Planning priority is not treated as a direct demand weight; it combines feasibility screening with priority-zone interpretation.

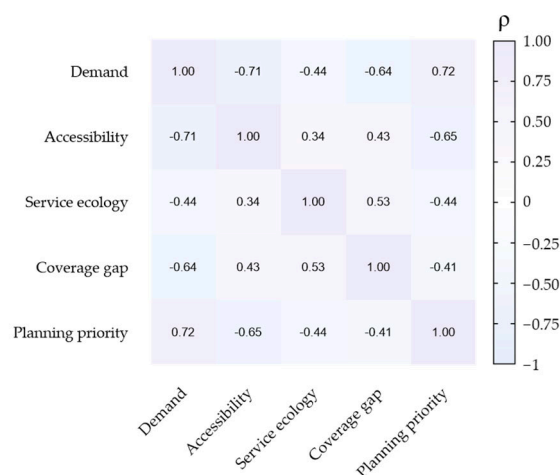


Figure 3. Dimension-level Spearman correlation and redundancy check.

4. Methods

4.1. Framework Overview

The proposed workflow transforms heterogeneous urban spatial datasets into a staged location-allocation optimisation problem. The framework first estimates observed clinic-location probabilities and constructs planning demand-weight proxies, then identifies candidate planning nodes and key priority areas, and finally solves a simultaneous priority-constrained covering model with a partitioned Stage 1/Stage 2 facility budget under alternative facility-budget scenarios, as shown in Figure 4.

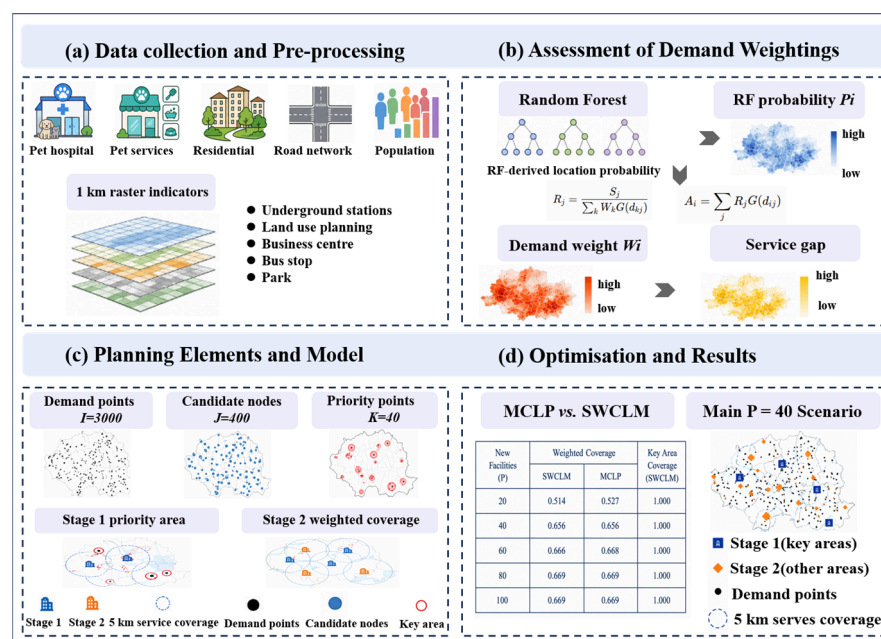


Figure 4. Stage-weighted covering location model framework.

4.2. Random Forest-Based Clinic-Location Probability and Planning Demand-Weight Construction

The RF classifier was used to characterize the spatial signature of existing veterinary hospital locations. Grid cells containing at least one validated veterinary hospital were defined as positive observations, while sampled grid cells without observed hospitals were treated as pseudo-negative observations. Because non-hospital cells may still reflect latent or unmet demand, the RF output is interpreted as an observed clinic-location prob-

ability or a proxy for location attractiveness, rather than as direct veterinary healthcare demand. Variables directly encoding existing hospital coverage, such as service-gap indicators and distance to the nearest hospital, were excluded from RF training to reduce the risk of target leakage.

The RF sampling design was specified as follows. The Beijing grid contains 16,828 cells, including 587 positive cells with 936 veterinary hospitals and 16,241 cells with no observed hospitals. Negative cells were sampled once at a 1:2 positive-to-negative ratio using a random seed of 42. All non-hospital grid cells were initially eligible for negative sampling, and no additional exclusion buffer around existing hospitals was imposed. Edge cells were retained when their centroids fell within the Beijing analysis boundary, whereas developability constraints were applied only during candidate-site screening rather than RF negative sampling. The model used an 80/20 stratified train–test split, 400 trees, `max_features = 'sqrt'`, `min_samples_leaf = 2`, `class_weight = 'balanced'`, bootstrap sampling, and 5-fold validation. No separate hyperparameter search was conducted; parameters not explicitly modified were retained at the scikit-learn defaults, including `max_depth = None` and `min_samples_split = 2`. Python 3.12, scikit-learn 1.9.0, and SciPy 1.17.1 were used for the analyses. Multiple hospitals within a single grid cell were treated as a single positive cell for classification, while hospital counts were retained as descriptive supply information.

To account for spatial dependence and the use of probability in the optimisation, this study reports both conventional random validation and district-grouped spatial validation. Additional class-imbalance and calibration metrics include balanced accuracy, specificity, the Matthews correlation coefficient, precision-recall AUC, Brier score, log loss, confusion matrix, and a calibration curve. The district-grouped pooled validation produced an ROC-AUC of 0.972 and an F1 score of 0.886, close to the random hold-out values, suggesting that the RF model is not solely an artifact of random spatial leakage. Nevertheless, the RF probability is retained strictly as a planning proxy and is not treated as externally validated service utilization.

Because adjacent grid cells are not statistically independent, the model interpretation focuses on predictive associations rather than causal effects. Previous studies using spatially structured validation have shown that RF testing splits may overestimate transferability [52]. In addition, RF variable-importance measures may be biased when predictors differ in scale or exhibit multicollinearity [53]. Therefore, this study reports both cross-validation performance and permutation importance, supplemented by VIF and indicator-correlation diagnostics.

The trained RF model produces a location probability for each grid cell. This probability is used as the observed-location component of the subsequent planning demand-weight proxy rather than as a direct measure of actual pet ownership, veterinary healthcare demand, or clinic utilization. To make the weight more suitable for allocation modeling, RF-derived probability was combined with population and residential-community intensity and then normalized across demand points. The SWCLM used the resulting weight as the relative demand value covered by candidate facilities. The demand weight was constructed as follows:

$$D_i = p_i \times (0.5 \text{Pop}_i^* + 0.5 \text{Res}_i^*) \quad (1)$$

$$w_i = D_i / \sum_i D_i \quad (2)$$

where p_i denotes the RF-derived observed clinic-location probability, Pop_i^* and Res_i^* represent normalized population and residential-community indicators, respectively, D_i denotes the unnormalized planning demand-weight proxy, and w_i denotes the final normalized demand weight. Accordingly, p_i is not interpreted as true demand. To reduce

over-interpretation and assess potential double-counting, the population-residential adjustment is reported as a transparent planning proxy and tested under alternative population/residential demand-weight specifications. Three constructions were compared: RF probability alone, population-residential index only, and the multiplicative combination of the RF probability and the population-residential index. The three variants all maintained complete key-area coverage at $P = 40$, but selected-node overlap with the main $P = 40$ solution ranged from 0.625 to 0.825. These results indicate that the multiplicative demand-weight construction affects the final allocation and should be interpreted as a planning proxy rather than as an empirically measured demand surface. A separate sensitivity analysis was also conducted for the population/residential coefficients. The tested 0.25/0.75, 0.40/0.60, 0.50/0.50, 0.60/0.40, and 0.75/0.25 settings produced high top-3000 demand overlap with the main surface (0.889–0.893) and very high Spearman rank correlation (approximately 0.992). In contrast, overlap among the top 40 priority areas ranged from 0.275 to 0.350 (Figure 5).

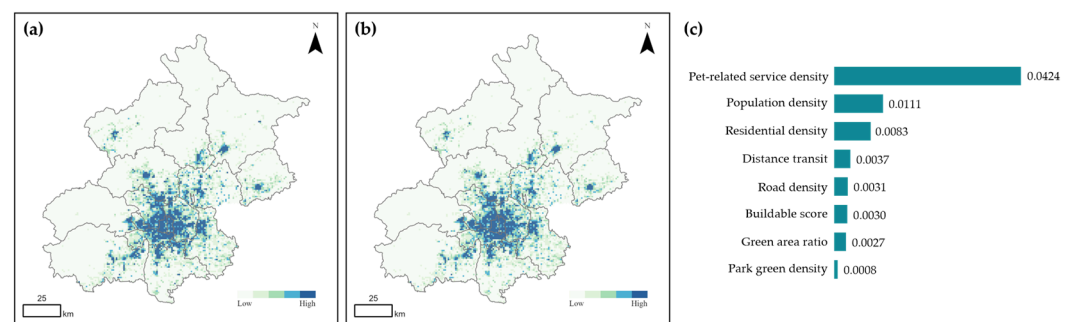


Figure 5. Random Forest-based clinic-location probability and planning demand-weight construction. The panels show (a) RF-derived location probability; (b) normalized demand weight and feature importance; and (c) hold-out classification performance.

4.3. Demand Points, Candidate Sites, and Key Priority Areas

The SWCLM is formulated using three spatial sets. The demand set I comprises 3000 high-demand grid centroids, the candidate set J comprises 400 eligible grid centroids screened by RF-derived location probability, service gap, demand intensity, buildability, and road accessibility, and the key set K comprises 40 priority areas with high demand weight, large service gaps, and weak existing coverage. The centroid coordinates serve only as computational representatives of the 1 km cells, not as verified parcels or entrances. Candidate eligibility therefore indicates grid-level planning feasibility; cadastral availability, building access, and street-entrance coordinates were unavailable and were not modeled. These set sizes retain citywide spatial detail while maintaining the tractability of mixed-integer optimisation.

The screening criteria and scoring rules are summarized in Table 5. Candidate sites are restricted to eligible grid cells with an area greater than 0.05 km² and a buildable score of at least 0.15, then ranked by a candidate score that combines observed clinic-location probability, E2SFCA service gap, demand-base score, buildability score, and road density. The associated coefficients are judgement-based planning weights rather than fitted parameters. Key priority areas are identified using a score combining the planning demand-weight proxy, E2SFCA service gap, and existing coverage gap. The reference priority set of $K = 40$ is defined independently of the tested facility's budget. A separate count-based sensitivity check varies the reference priority-set size across $K = 20, 30, 40, 50$, and 60 at $P = 40$ under the same 5 km catchment. In addition, the revision systematically perturbs every candidate-score and key-area-score coefficient by $\pm 20\%$, renormalizes each perturbed weight vector to sum to one, and tests buildability cutoffs of 0.10, 0.15, and 0.20. A separate

threshold-derived priority set is defined before facility-budget testing by selecting eligible cells at or above the 99th percentile of the key-area score (Figure 6).

Table 5. Candidate and key-priority-area selection rules.

Item	Rule or Value	Rationale
Candidate eligibility	eligible for new site = 1; area > 0.05 km ² , buildable score ≥ 0.15	Removes water/protected/very small cells before optimisation
Candidate score	0.38 RF probability + 0.24 E2SFCA service gap + 0.18 demand base + 0.12 buildability + 0.08 road density	Balances observed market signal, underserved demand, feasibility, and access
Candidate set	Top 400 eligible grid centroids, score range 0.763–0.912	Keeps citywide candidate choice while maintaining MILP tractability
Key-area score	0.45 planning demand-weight proxy + 0.35 E2SFCA service gap + 0.20 existing coverage gap	Defines priority areas as high-demand and underserved cells
Key priority areas	Top 40 eligible populated grid cells ranked by the key-area score; score range 0.703–0.820	Reference priority-set size used independently of facility budget; K = 20/30/40/50/60 sensitivity was used to test dependence on this choice

Note: All score components were min-max normalized before weighted summation.

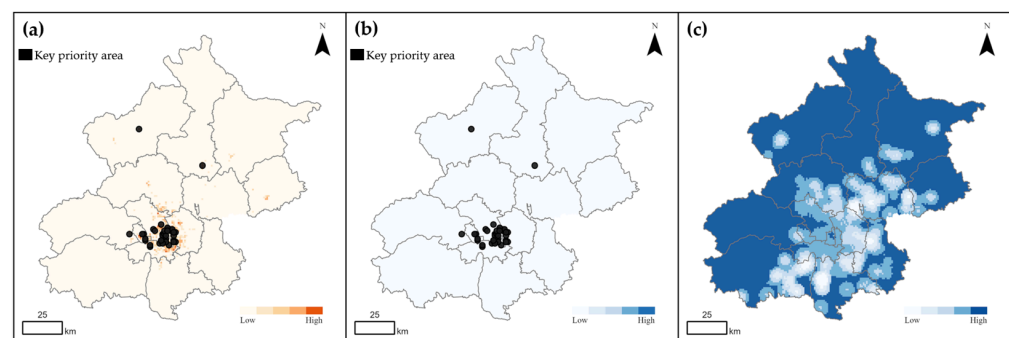


Figure 6. SWCLM input sets. The panels show (a) candidate planning nodes, (b) key priority areas, and (c) demand-point context over the existing service-gap surface.

4.4. Stage-Weighted Covering Location Model

The SWCLM selects candidate sites subject to a fixed facility-budget constraint. Let I , J , and K denote the sets of demand points, candidate sites, and key priority areas, respectively. Each demand point i is assigned a normalized weight w_i . The binary parameter a_{ij} equals one if candidate site j covers demand point i within the service catchment, and zero otherwise. Similarly, the binary parameter b_{jk} equals one if candidate site j covers key priority area k . The decision variables x_{j1} and x_{j2} indicate whether candidate site j is selected in Stage 1 or Stage 2, respectively, while $y(i)$ indicates whether demand point i is covered by at least one selected site. The parameters and their meanings are shown in Table 6.

Table 6. Parameters and meanings.

Parameter	Meaning
I	Demand points
J	Candidate planning sites

Table 6. Cont.

Parameter	Meaning
K	Key priority areas
i	Index of demand points, $i \in I$
j	Index of candidate planning sites, $j \in J$
k	Index of key priority areas, $k \in K$
p_i	RF-derived observed clinic-location probability for demand point i
a_{ij}	Binary coverage parameter equal to 1 if $d_{ij} \leq R$; otherwise, it is 0
x_{j1}	Binary decision variable equal to 1 if candidate site j is selected in Stage 1; otherwise, it is 0
x_{j2}	Binary decision variable equal to 1 if candidate site j is selected in Stage 2; otherwise, it is 0
y_i	Binary decision variable equal to 1 if demand point i is covered by at least one selected site; otherwise, it is 0
d_{jk}	Distance between candidate site j and key priority area k
b_{jk}	Binary coverage parameter equal to 1 if $d_{jk} \leq R$; otherwise, it is 0

Stage 1 variables enforce coverage of all key priority areas, while Stage 2 variables use the remaining facility budget to improve aggregate weighted coverage. Both sets of variables are optimised simultaneously within the same mixed-integer linear programming (MILP) formulation. The coverage parameters are defined by a service radius R :

$$a_{ij} = 1 \text{ if } d_{ij} \leq R; \text{ otherwise } a_{ij} = 0 \quad (3)$$

$$b_{jk} = 1 \text{ if } d_{jk} \leq R; \text{ otherwise } b_{jk} = 0 \quad (4)$$

The SWCLM objective function and main constraints are expressed as follows:

The implemented model is formulated as a simultaneous priority-constrained covering model with a partitioned facility budget rather than a fully sequential two-stage optimisation solved in two independent passes. Stage 1 variables are constrained to ensure coverage of all predefined priority areas, while Stage 2 variables are optimised to increase aggregate weighted coverage within the remaining budget. The term stage-weighted is therefore used in the inherited sense of stage-prioritised and budget-partitioned coverage, not as a claim that separate stage weights enter the objective function. The notation table and binary-domain expression were used to indicate that x_{j1} , x_{j2} , and y_i belong to $\{0, 1\}$. As a computational relaxation, y_i may equivalently be treated as a continuous variable on $[0, 1]$ without changing the optimal solution because the objective maximises covered demand and the coverage constraint upper-bounds y_i by selected facilities.

The E2SFCA service-gap indicator was calculated using existing veterinary hospitals as equal-capacity point supplies. Existing veterinary hospitals were treated as equal-capacity point supplies because verified information on clinic floor area, veterinarians, equipment, opening hours, and consultation volume was unavailable. For supply site j , the supply-to-demand ratio was calculated as $R_j = \sum_k D_k W_{kj}$, where $S_j = 1$, D_k denotes the demand proxy for grid cell k , and W_{kj} denotes the distance-decay weight. The implemented distance-decay weights were 1.00 for 0–1 km, 0.65 for 1–3 km, and 0.30 for 3–5 km. Grid accessibility was then calculated as $A_i = \sum_j R_j W_{ij}$ for hospitals j within the specified catchment of demand grid i . The service-gap indicator was defined as the min-max-

normalized inverse of A_i ; grid cells with zero demand were excluded from demand-point selection, and cross-boundary demand or supply outside the Beijing analysis boundary was not included. The main optimisation adopted a 5 km catchment, while the 3 km radius was retained only as a descriptive proxy for existing coverage. This equal-capacity assumption is identified as a major limitation, and the model is described as estimating planning-priority coverage rather than clinic-level service capacity.

$$\text{Maximize } Z = \sum_i \in I w_i y_i \quad (5)$$

$$y_i y_i \leq \sum_j \in J a_{ij} (x_{j1} + x_{j2}), \forall i \in I \quad (6)$$

$$\sum_j \in J b_{jk} x_{j1} \geq 1, \forall k \in K \quad (7)$$

$$\sum_j \in J x_{j1} = P_1; \sum_j \in J x_{j2} = P_2; x_{j1} + x_{j2} \leq 1, \forall j \in J \quad (8)$$

$$x_{j1}, x_{j2}, y_i \in \{0, 1\}, \forall j \in J, \forall i \in I \quad (9)$$

P_1 and P_2 denote the facility budgets for Stages 1 and 2, respectively, and the total budget P is defined as $P = P_1 + P_2$. The resulting formulation is a MILP problem that can be solved with standard MILP solvers such as Gurobi, CPLEX, or HiGHS. In this study, the optimisation model was implemented in Python and solved using HiGHS via `scipy.optimize.milp`. The primary analysis adopted a 5 km service range, whereas descriptive mapping of existing coverage uses a 3 km service coverage range as a proxy indicator. This distinction is consistent with the general principles of spatial healthcare accessibility analysis and floating service range methods [45–48]. Because the selected nodes represent strategic planning locations rather than parcel-level or final commercial addresses, the resulting allocations are interpreted as planning-priority nodes. Existing hospitals influence priority identification through the E2SFCA-derived service-gap and descriptive coverage-gap surfaces, but they are not fixed facility variables in the MILP coverage constraints. Consequently, the reported weighted coverage represents coverage provided by the proposed planning nodes conditional on the existing-service-gap surface, not total citywide combined coverage by existing and proposed hospitals. A Stage 1 feasibility assessment showed that all 40 predefined priority areas had at least one feasible candidate node within the 5 km catchment, with candidate multiplicity ranging from 1 to 57. If the Stage 1 budget P_1 is smaller than the minimum number of candidate facilities required to cover all priority areas in K , constraints (7) and (8) render the model infeasible. In such cases, P_1 , the definition of the priority set, or the service catchment must be adjusted before optimisation.

The accessibility component is also consistent with later refinements of floating-catchment analysis. Modified and variable distance-decay formulations have shown that provider competition, catchment size, and distance-decay assumptions can substantially affect local access estimates [54,55], while supply–demand adjustment explicitly addresses competition for limited service capacity [56]. In this study, these methods provide the basis for diagnosing spatial service gaps. However, they do not convert commercial veterinary hospital systems into public healthcare systems with known service capacities.

4.5. Baseline and Scenario Design

The SWCLM was evaluated against the MCLP baseline. The MCLP baseline uses the same demand points, candidate sites, demand weights, and 5 km catchment, but excludes the Stage 1 key-area coverage constraint. This comparison provides a controlled evaluation because both models are based on identical spatial inputs and facility budgets. The tested facility budgets are 20, 40, 60, 80, and 100 new facilities. In the SWCLM, 60% of each budget was allocated to Stage 1 priority-area coverage and 40% to Stage 2 weighted

demand coverage. This ratio reflects the intended prioritisation of function over aggregate weighted coverage, which is consistent with the purpose of the phased model. A sensitivity check under the 40-facility scenario compares 50/50, 60/40, and 70/30 stage allocations. All three ratios achieve complete key-area coverage and the same weighted coverage to three decimal places, indicating that the specific 60/40 split did not drive the main conclusion. The evaluation metrics include weighted demand coverage, demand-point coverage rate, and key-area coverage rate, as shown in Table 7.

Table 7. Stage-allocation ratio sensitivity under the 40-facility scenario.

Stage 1 Share	Stage 1 Sites	Stage 2 Sites	Weighted Coverage	Demand-Point Coverage	Key-Area Coverage
50%	20	20	0.656	0.562	1.000
60%	24	16	0.656	0.562	1.000
70%	28	12	0.656	0.562	1.000

Note: The 60/40 setting is retained because it protects key areas while preserving the same weighted coverage as adjacent 50/50 and 70/30 settings.

The 3 km radius was used only as a descriptive proxy for existing service coverage, whereas the 5 km radius defines the main optimisation catchment. To examine the influence of this assumption, the $P = 40$ scenario was evaluated for catchments of 3, 5, 7, and 10 km. Weighted coverage increased from 0.455 at 3 km to 0.656 at 5 km, 0.743 at 7 km, and 0.825 at 10 km, while key-area coverage remained 1.000 in all tested radii. Accordingly, the 5 km radius is interpreted as a modeling scenario rather than an empirically established veterinary travel threshold.

Five additional sensitivity analyses were conducted under the 5 km catchment. First, a single-factor sensitivity analysis was performed on the baseline candidate weights (0.38, 0.24, 0.18, 0.12, and 0.08) and key-area weights (0.45, 0.35, and 0.20), perturbing each coefficient by $\pm 20\%$, and the corresponding weight vector was renormalized. Candidate-set overlap, priority-set overlap, selected-node overlap, weighted coverage, and priority-area coverage were recorded at $P = 40$. Second, the reference priority-set size was varied across $K = 20, 30, 40, 50$, and 60 at $P = 40$ while holding the facility budget and 5 km catchment fixed; weighted coverage, demand-point coverage, key-area coverage, and solver status were recorded. Third, threshold-derived priority sets were generated using the 98.5th, 99th, and 99.5th percentiles of the key-area score; the prespecified 99th-percentile set was then evaluated at $P = 20, 40$, and 60. Fourth, isotonic RF calibration was performed only within the training sample using 3-fold internal cross-validation. The raw and calibrated probability surfaces were each propagated through demand weighting, candidate screening, priority-set construction, and the $P = 40$ allocation model. Fifth, candidate-centroid positional sensitivity was evaluated at $P = 40$ by independently relocating all 400 candidate coordinates to random points within the intersection of their original grid polygon and 250 m or 500 m centroid buffers, or the entire original 1 km grid. Ten fixed-seed replicates were run per setting (30 perturbed models); demand and key-area coordinates were held constant, each coverage matrix was rebuilt, and the MILP model was re-solved for each replicate. These random positions represent within-cell positional uncertainty rather than verified feasible parcels.

MCLP remains the main benchmark because it uses the same candidate set, demand weights, and catchment radius without the key-area assurance constraint. The proposed model is not claimed to be universally superior; instead, it exposes the trade-off between aggregate weighted coverage and guaranteed priority-area coverage. The advantage is expected to be most relevant under lower facility budgets, where MCLP leaves some key areas uncovered.

4.6. Multicollinearity Diagnostics

Because several urban-service indicators may exhibit spatial correlation, multicollinearity was assessed before interpreting RF feature importance. Variance inflation factors (VIFs) were calculated for all RF input variables. Most variables exhibited VIF values below 2.5, indicating limited multicollinearity among the predictors. The highest VIF values were observed for the buildability score and green-area ratio, which is expected because both variables reflect related aspects of land-use constraints and spatial feasibility. Variables with VIF values below 10 were retained because they represent planning feasibility rather than the same demand mechanisms captured by population density, residential density, or pet-related service density. Figure 3 and Table 8 jointly indicate that these indicators contain correlated yet non-identical spatial information.

Table 8. Multicollinearity diagnostics for RF input variables.

Variable	VIF	Interpretation
Buildable Score	5.65	Moderate
Green Area Ratio	5.60	Moderate
Road Density	2.45	Low
Population Density	1.75	Low
Distance to Major Road	1.52	Low
Residential Density	1.52	Low
Pet-related Service Density	1.47	Low
Distance to Transit	1.38	Low
Commercial Density	1.34	Low
Park Green Density	1.15	Low

Note: VIF values below 10 are generally not treated as severe multicollinearity; the two moderate values are land-constraint variables and are interpreted cautiously.

5. Results

5.1. Existing Veterinary Hospital Distribution and Input Spatial Patterns

The POI dataset contains 936 veterinary hospitals in Beijing, with a highly uneven spatial distribution across districts. Chaoyang District has the largest number of veterinary hospitals (221 facilities), followed by Fengtai, Changping, Daxing, Haidian, Tongzhou, Fangshan, and Shunyi. Peripheral districts contain substantially fewer facilities, reflecting the combined influence of lower population density, weaker commercial activity, and greater ecological and mountainous constraints. This existing distribution provides the observed location pattern that the RF model learns from, but it should not be interpreted as an optimal service configuration.

The input covariate maps show that demand-related, accessibility-related, and service-gap indicators are spatially correlated but not identical. Population density and residential density are primarily concentrated in central and inner-suburban areas of Beijing, whereas road density follows major development corridors. Pet-related service density and commercial density highlight mixed residential-commercial clusters, while the service-gap index identifies areas where high demand potential coincides with limited existing accessibility. These spatial mismatches among covariates highlight the rationale for a multi-step allocation framework rather than relying solely on descriptive spatial mapping.

5.2. Random Forest Location-Probability and Planning-Weight Estimation

The RF model shows strong predictive performance for classifying observed veterinary hospital location patterns, although its output remains a planning proxy. In the random hold-out set, accuracy was 0.926, balanced accuracy 0.924, precision 0.871, recall 0.915, specificity 0.932, F1 0.893, MCC 0.837, ROC-AUC 0.971, precision-recall AUC 0.933, Brier

score 0.062, and log loss 0.211. The district-grouped spatial validation produced a pooled accuracy of 0.920, an F1 score of 0.886, an ROC-AUC of 0.972, and a PR-AUC of 0.940. These results indicate a strong association between urban covariates and observed clinic-location patterns, but they do not constitute validation of actual pet ownership, expenditure, consultation volume, or unmet healthcare demand (Table 9).

Table 9. Random Forest classification performance.

Metric	Value
Positive prevalence	0.334
Accuracy	0.926
Balanced accuracy	0.924
Precision	0.871
Recall	0.915
Specificity	0.932
F1 score	0.893
MCC	0.837
ROC-AUC	0.971
PR-AUC	0.933
Brier score	0.062
Log loss	0.211
ECE, quantile-bin approximation	0.040
Confusion matrix (TN/FP/FN/TP)	219/16/10/108

Note: Values are rounded to three decimals.

The completed calibration experiment showed mixed rather than uniformly improved predictive calibration. On the unchanged hold-out set, the raw RF and isotonic-calibrated RF model achieved ROC-AUC values of 0.971 and 0.969, Brier scores of 0.0621 and 0.0614, log losses of 0.211 and 0.229, and 10-bin ECE values of 0.040 and 0.036, respectively. Thus, isotonic calibration slightly improved Brier score and ECE, but worsened log loss and marginally reduced discrimination. After each probability surface was propagated through the complete allocation workflow, both $P = 40$ solutions covered all priority areas. The raw and calibrated weighted coverage values were 0.656 and 0.709, respectively, and the corresponding demand-point coverage values were 0.562 and 0.583, candidate-set overlap was 93.0%, and selected-node overlap between the two solutions was 57.5%. Because the two solutions use differently calibrated and renormalized demand surfaces, the higher calibrated weighted-coverage value is not interpreted as direct performance superiority; the node overlap demonstrates that the choice of calibration materially affects local recommendations.

A pet-service ablation analysis was added because pet-service density may partly reflect spatial agglomeration around existing veterinary hospitals. Removing pet-service density reduced ROC-AUC from 0.971 to 0.955 and F1 from 0.893 to 0.855 in the random hold-out set. This indicates that pet-service density provides additional predictive information, but its interpretation is now limited to spatial association with the pet-service ecosystem rather than causal demand generation (Table 10).

Feature-importance analysis shows that pet-related service density is the most important predictor according to permutation importance, followed by population density and residential density. Transit distance, road density, and buildability also contribute to model performance, whereas commercial density has a weaker independent effect after the other variables are included. This pattern suggests that veterinary hospital locations are more closely associated with residential and pet-service ecosystems than with general commercial activity alone. From a planning perspective, areas with dense pet-related ser-

vice ecosystems and high residential density require particular attention, even when they are not major commercial centers.

Table 10. Leading RF permutation-importance variables.

Feature	Permutation Importance	SD
Pet-related Service Density	0.0424	0.0120
Population Density	0.0111	0.0031
Residential Density	0.0083	0.0039
Distance to Transit	0.0037	0.0028
Road Density	0.0031	0.0017
Buildable Score	0.0030	0.0024
Green Area Ratio	0.0027	0.0022
Park Green Density	0.0008	0.0010

Note: Permutation importance measures the reduction in predictive performance when each feature is perturbed.

5.3. Candidate Sites and Key Priority Areas

The proposed nodes are primarily concentrated in developed and accessible areas of Beijing, while also including selected suburban locations with substantial service gaps and high potential demand. This pattern reflects the screening criteria: candidate planning nodes must be feasible for facility planning rather than simply correspond to areas with the highest existing veterinary hospital density. The predefined priority areas constitute a more selective subset, highlighting locations with relatively high planning-demand weights and limited existing service coverage that receive guaranteed consideration in Stage 1. This distinction between candidate planning nodes and key priority areas is central to the SWCLM framework, because it prevents the model from allocating all resources to the highest aggregate-demand locations while overlooking strategically underserved areas.

5.4. Facility-Budget Scenarios and Baseline Comparison

The SWCLM scenario results reveal a clear pattern of diminishing returns. With 20 proposed facilities, the model achieves weighted coverage of 0.514 and demand-point coverage of 39.2% while ensuring complete key-area coverage. Increasing the facility budget to 40 raises weighted coverage to 0.656 and demand-point coverage to 56.2%. Further increases to 60, 80, and 100 facilities produce additional but smaller gains, with weighted coverage stabilizing at approximately 0.669. Therefore, the 40-facility scenario does not achieve the numerical maximum; it represents an elbow-type planning scenario that captures most of the attainable coverage under the tested assumptions, as shown in Figure 7 and Table 11.

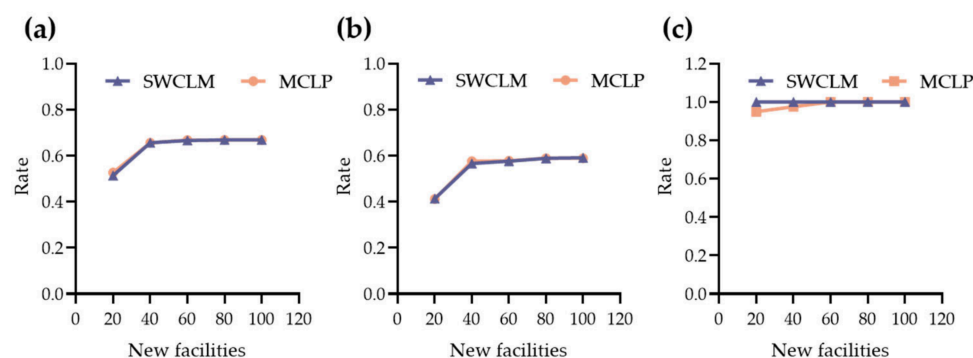


Figure 7. Coverage rates are compared under 20, 40, 60, 80, and 100 proposed new facilities. (a) Weighted demand coverage, (b) demand-point coverage, and (c) key-area coverage.

Table 11. SWCLM and MCLP scenario comparison.

New Facilities	Stage 1	Stage 2	SWCLM Weighted	MCLP Weighted	SWCLM Key-Area	MCLP Key-Area
20	12	8	0.514	0.527	1.000	0.950
40	24	16	0.656	0.656	1.000	0.975
60	36	24	0.666	0.668	1.000	1.000
80	48	32	0.669	0.669	1.000	1.000
100	60	40	0.669	0.669	1.000	1.000

Note: Coverage values range from 0 to 1. SWCLM includes the Stage 1 key-area assurance constraint; MCLP optimises aggregate weighted coverage only.

The coverage-rate curves and numerical differences in key-area coverage should be interpreted together. The MCLP baseline achieves similarly or slightly higher overall weighted demand coverage because its objective exclusively maximises aggregate demand coverage. However, under resource-constrained conditions, it does not adequately protect all key priority areas. With 20 facilities, the MCLP achieves a weighted demand coverage of 0.527 and covers 95.0% of key priority areas; with 40 facilities, it achieves a weighted demand coverage of 0.656 and covers 97.5% of key priority areas. By contrast, the SWCLM guarantees 100% key-area coverage in all tested scenarios. These results indicate that the SWCLM is not intended to optimise all performance metrics simultaneously; instead, it explicitly balances aggregate coverage efficiency against guaranteed priority-area coverage. In absolute terms, this means that SWCLM increases coverage in key priority areas by two additional areas at $P = 20$ (40 rather than 38 of 40) and by one additional key priority area at $P = 40$ (40 rather than 39 of 40). A demand-ranked heuristic baseline was also added as a non-MILP comparator; at $P = 40$ it covered only 0.314 of weighted demand and 92.5% of key priority areas, indicating that explicit covering optimisation substantially outperforms simple selection of the highest-ranked candidate nodes in this scenario. Because SWCLM, MCLP, and the heuristic are deterministic solutions evaluated on the same fixed spatial instance, conventional inferential significance testing of a single paired solution is not directly applicable; therefore, robustness is assessed through scenario and sensitivity analyses rather than through p -values (Table 12).

Table 12. Robustness and sensitivity analyses of key modeling assumptions.

Concern	Tested Setting	Main Result	Interpretation
Candidate set size	J = 200/400/600 at $P = 40$	J = 200: 0.525 weighted coverage; J = 400: 0.656; J = 600 reached solver time limit; J = 800/all-feasible were not completed after the J = 600 time limit	The 400-candidate case achieves substantially higher weighted coverage than J = 200, whereas larger candidate sets require greater computational effort.
Demand double-counting	RF only; Pop-Res only; RF x Pop-Res	All variants kept key-area coverage = 1.000; selected-node overlap = 0.625–0.825	The demand construction affects allocation and is treated as a planning proxy.
RF calibration and allocation	Raw RF versus training-only 3-fold isotonic calibration at $P = 40$	Brier: 0.0621 versus 0.0614; ECE: 0.040 versus 0.036; weighted coverage: 0.656 versus 0.709; selected-node overlap: 0.575	Calibration modestly improves two calibration metrics but materially changes local allocations; no universal improvement is claimed.

Table 12. Cont.

Concern	Tested Setting	Main Result	Interpretation
Candidate/key-score coefficients	All coefficients $\pm 20\%$, renormalized; buildability cutoff 0.10/0.15/0.20; 19 variants at $P = 40$	Weighted coverage 0.635–0.682; priority coverage 1.000; candidate overlap ≥ 0.973 ; key-set overlap ≥ 0.900 ; selected-node overlap 0.575–0.775	Aggregate coverage and priority assurance are stable, while the exact selected grids vary across plausible parameter settings.
Threshold-derived priority set	98.5th/99th/99.5th percentiles; 99th-percentile set tested at $P = 20/40/60$	99th percentile generated 169 cells; weighted coverage 0.476/0.654/0.662 and priority coverage 1.000 at $P = 20/40/60$; 98.5th percentile (253 cells) was infeasible at $P = 40$	Priority need is defined independently of facility budget, and the wider threshold identifies a Stage 1 capacity boundary.
Key-area count	$K = 20/30/40/50/60$ at $P = 40$	All five runs optimal; weighted coverage 0.6539–0.6543; demand-point coverage 0.5587–0.5597; key-area coverage 1.000	Aggregate coverage is highly stable across the tested reference priority-set sizes; exact site-level invariance is not claimed.
Candidate-centroid position	Candidate coordinates perturbed within 250 m, 500 m, or the original 1 km grid; 10 seeds per setting at $P = 40$	30/30 runs optimal; weighted coverage 0.6541–0.6659; demand-point coverage 0.5597–0.5743; key-area coverage 1.000; exact-node overlap 0.525–0.775	Aggregate performance and priority assurance are stable, but exact nodes are sensitive to within-cell position; outputs are grid-level search zones.

Across the 19 coefficient-perturbation and buildability-threshold variants, all $P = 40$ models were successfully solved and retained complete priority-area coverage. Weighted coverage ranged from 0.635 to 0.682, candidate-set overlap with the baseline solution remained above 97.25%, and key-set overlap was at least 90.0%. Selected-node overlap with the reference solution $P = 40$ ranged from 57.5% to 77.5% (median 72.5%), indicating that overall service performance is stable, but the exact spatial configuration of selected nodes varied across plausible parameter settings. Across the separate key-area-count sensitivity runs at $P = 40$, $K = 20, 30, 40, 50$, and 60, all scenarios reached optimal solutions and achieved complete priority-area coverage. Weighted coverage varied only from 0.6539 to 0.6543 and demand-point coverage from 0.5587 to 0.5597; these results indicate that aggregate coverage was highly stable across the tested reference priority-set sizes; identical site-level solutions cannot be expected. The independently defined 99th-percentile priority rule generated 169 cells and achieved complete priority coverage at $P = 20, 40$, and 60, with weighted coverage of 0.476, 0.654, and 0.662, respectively. The 99.5th percentile set contained 85 cells and was feasible at $P = 40$, whereas the broader 98.5th-percentile set contained 253 cells and made the $P = 40$ Stage 1 constraint infeasible. These findings provide a budget-independent robustness check beyond the fixed top-40 reference definition and reveal the feasibility boundary of the priority constraint. A final sensitivity experiment examined candidate-centroid displacement. Across 30 perturbed $P = 40$ models (10 replicates for each spatial perturbation setting), all runs reached optimal solutions and retained 100% key-area coverage. Weighted coverage ranged from 0.6541 to 0.6659 and demand-point coverage from 0.5597 to 0.5743. Selected-node overlap with the unperturbed rerun ranged from 52.5% to 77.5% (mean = 65.8%), while 9.7–13.1% of candidate-demand coverage links changed. Therefore, aggregate coverage and priority assurance were stable, but within-cell positional uncertainty materially affected the exact node configuration.

5.5. Spatial Allocation Results and the Main $P = 40$ Scenario

The results indicate the spatial evolution of the SWCLM allocations under different facility-budget scenarios. Under $P = 20$, selected nodes are primarily concentrated in key priority areas and adjacent high planning-demand clusters, reflecting the model's emphasis on ensuring key priority-area coverage under resource constraints. Increasing the budget to $P = 40$ expands the allocation across both central and suburban areas of Beijing, with additional Stage 2 nodes extending weighted-demand coverage while maintaining complete coverage of key priority areas. As the budget increases further to $P = 60$ and $P = 100$, additional nodes are gradually allocated among the remaining demand clusters; however, the overall spatial configuration becomes progressively more stable. This pattern indicates diminishing marginal spatial expansion beyond $P = 40$, suggesting that the 40-facility scenario captures the major priority and high-demand areas under the tested assumptions rather than representing a strict numerical optimum, as shown in Figure 8.

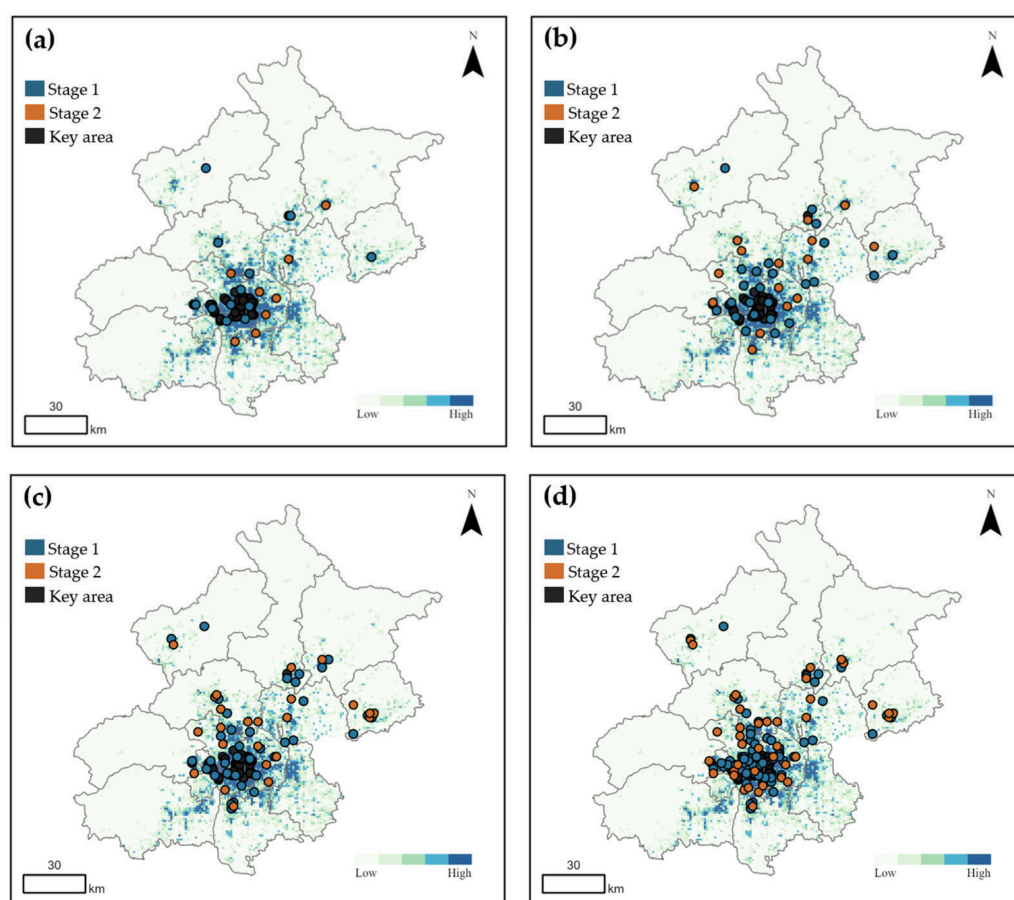


Figure 8. SWCLM allocation maps under different facility budgets. (a) $P = 20$, (b) $P = 40$, (c) $P = 60$, (d) $P = 100$.

The $P = 40$ scenario is interpreted as the primary planning scenario rather than a formally optimal facility number. Under this scenario, 24 Stage 1 sites are allocated to ensure priority-area coverage, and 16 Stage 2 sites are assigned to improve weighted-demand coverage. The model achieves weighted coverage of 0.656, demand-point coverage of 0.562, and complete key-area coverage. Because reliable information on facility costs, licensing requirements, rental conditions, staffing availability, and service capacity is unavailable, this manuscript now avoids recommending exactly 40 immediately implementable hospitals. Instead, $P = 40$ is presented as a useful scenario for demonstrating the marginal-coverage elbow and the spatial implications of priority-constrained allocation, as shown in Figure 9.

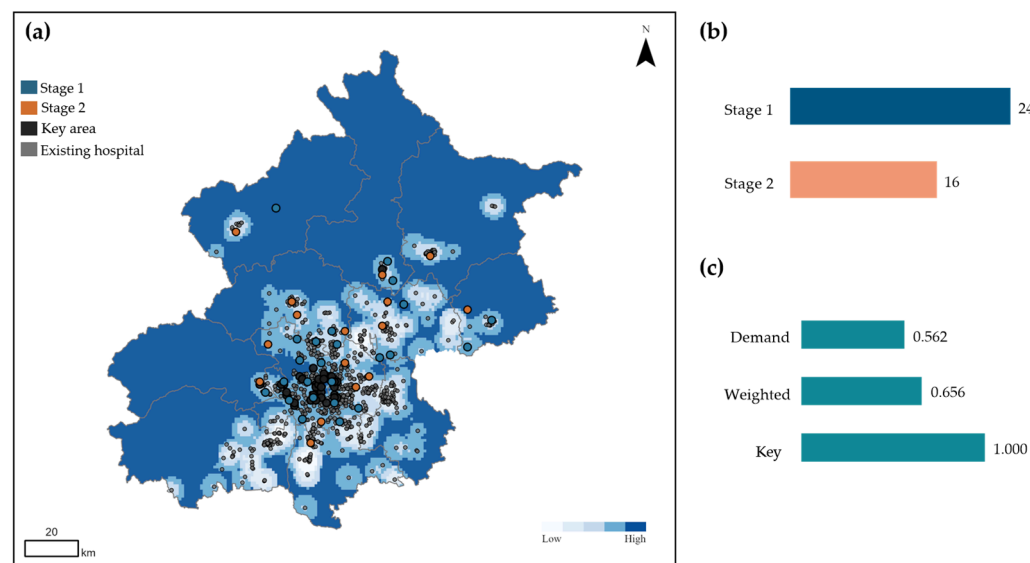


Figure 9. Main $P = 40$ scenario SWCLM allocation over the service-gap surface. (a) SWCLM allocation over service gap ($P = 40$), (b) stage composition, (c) coverage performance.

6. Discussion

6.1. From Suitability Mapping to Allocation Optimisation

The main contribution of this study is not a new universally applicable optimisation algorithm, but a transparent adaptation of staged covering logic to the metropolitan veterinary healthcare planning problem. This contribution can be understood at three complementary levels: methodological adaptation, including RF-based clinic-location probability, E2SFCA service-gap screening, and priority-constrained budget partitioning; domain-specific methodological reconstruction, including extending a covering framework originally developed for bounded scenic-area waste-bin allocation to open urban veterinary healthcare; and application contribution, including Beijing-wide scenario evidence for planning-priority nodes.

6.2. Planning Implications for Beijing

The results indicate that veterinary hospital planning in Beijing cannot be understood simply as a shortage problem in the urban core. The urban core already contains a large number of veterinary hospitals and is characterized by high population density, intensive residential development, and a well-established pet-related service ecosystem. Peripheral districts generally have fewer veterinary hospitals and larger service gaps, but some of these areas also face constraints related to low buildability or weaker demand-related planning indicators. Therefore, effective allocation requires differentiating among areas with high demand potential and strong existing competition, areas with strong demand-related signals but inadequate service coverage, and extensive low-density areas where new facilities may not be immediately viable. The locations identified by the SWCLM provide a spatial representation of these differentiated planning priorities. This interpretation aligns with recent accessibility studies, which emphasize service accessibility and inclusive urban access as planning objectives rather than relying solely on facility counts [31,32].

6.3. Interpretation of Candidate Planning Nodes

Selected nodes are grid-level planning locations, not final commercial addresses. Each is represented by the centroid of a 1 km cell, which may fall on an unavailable or inaccessible site; the nearest feasible parcel or street entrance may lie near the cell edge or in an adjacent cell. Such displacement can change road access, competition, and 5 km catch-

ment membership, thereby altering candidate rankings and exact selected nodes, especially near catchment boundaries. The outputs therefore identify search zones, not build-ready parcels. Implementation requires parcel-level verification of land use, vacancy, rent, ownership, licensing, parking, access, and network travel time, followed where necessary by re-optimisation with verified parcel or entrance coordinates.

6.4. Relationship to the Stage-Covering Reference Model

The relationship between this study and the Yunshan waste-bin study demonstrates both methodological continuity and domain-specific reformulation. Both studies adopt a staged coverage framework that ensures key areas are protected before optimising weighted demand coverage; however, their decision contexts differ substantially. The scenic-area model addresses standardized public sanitation facilities within a bounded visitor system. In contrast, the Beijing case concerns private pet healthcare services in an open metropolitan market characterized by overlapping service areas and heterogeneous land conditions. Accordingly, this study replaces waste-generation intensity with an RF-derived observed clinic-location probability incorporated into a planning demand-weight proxy; incorporates residential and pet-service ecology; embeds E2SFCA-based accessibility gaps into priority screening; and interprets selected nodes as strategic planning zones rather than directly deployable facility locations. Previous waste-location studies demonstrate the value of covering models, multicriteria trade-offs, and integrated allocation approaches [57–60], while Digital Earth and spatial-optimisation studies support the broader transition from descriptive geospatial analysis toward decision-oriented models [61–68]. Therefore, the contribution of this study is not a simple relabeling or application of the SWCLM, but a transparent reformulation of demand representation, competitive context, and feasibility assessment for metropolitan pet healthcare planning.

This interpretation also clarifies that the framework may transfer to other cities only after local recalibration. Compact central cities, polycentric metropolitan regions, and rapidly suburbanizing cities may require different candidate-screening weights, catchment radii, capacity proxies, and priority-area criteria. Therefore, the Beijing results should be read as a city-specific strategic allocation case rather than as external validation of a universal veterinary facility model.

6.5. Limitations and Future Work

Several limitations of this study should be acknowledged. First, the RF component learns spatial patterns from existing veterinary hospital locations and may reproduce historical market concentration, rental-price effects, regulatory constraints, and agglomeration patterns associated with pet-service clusters, and RF uncertainty is not fully propagated through the MILP. Future work should use bootstrapped probability surfaces to identify selection frequency and robust candidate nodes. Second, the integration of population data from 2020 with POI/OSM data collected in 2026 introduces temporal inconsistency; contemporary residential POIs partly update the spatial signal, but cannot fully replace recent population or pet-ownership data; the manual validation of 30 POI records is treated as a preliminary spot-check rather than comprehensive quality validation, and a larger district-stratified validation sample is recommended. Third, Euclidean distance and equal clinic capacity are planning proxies that should be replaced by network travel time and verified capacity attributes where such data are available. The fixed 1 km grid introduces both modifiable areal unit problem (MAUP) effects and within-cell positional uncertainty: alternative cell sizes, zoning systems, or substitution of a centroid with a feasible parcel or entrance could change catchment membership, local rankings, and exact node selection. Candidate-score and key-area-score coefficients remain judgement-based plan-

ning weights. The completed perturbation experiment shows stable aggregate coverage and priority assurance, but only moderate exact-node overlap, so selected cells should be interpreted as planning alternatives rather than unique sites. Future applications should construct parcel-level candidate sets, rerun the allocation with verified parcel or entrance coordinates, and report node-selection stability after coordinate substitution. Future work should improve empirical demand measurement and strengthen external validation. Recent studies of veterinary accessibility have more directly estimated community pet populations, clinic availability, access barriers, and veterinary service deserts [69,70]. If comparable data for Beijing become available, including pet-owning households, consultation volumes, staffing, opening hours, prices, and clinic capacity, the RF-based demand proxy and equal-capacity assumption could be replaced with more direct demand and supply measures. The identification of key priority areas should also be validated using road-network travel time, field investigations, expert assessment, and subsequent clinic-opening outcomes to determine whether the modeled priorities correspond to actual improvements in accessibility and commercial feasibility.

7. Conclusions

This study applies and adapts a priority-constrained covering allocation framework for veterinary hospital planning in Beijing. Using multi-source POI, population, road network, green space, and administrative boundary data, this study develops 1 km grid-level spatial covariates, estimates RF-based observed clinic-location probability, generates candidate planning nodes and key priority areas, and evaluates facility-budget scenarios. The results show that pet-related service density, population density, and residential-community density are important predictors of existing clinic-location patterns. Under the primary 5 km catchment assumption, the $P = 40$ scenario achieves complete priority-area coverage and captures most of the attainable weighted coverage; however, it should be interpreted as an elbow-type planning scenario rather than a demonstrated commercial optimum. The framework supports strategic planning priority analysis, while final hospital siting requires further consideration of parcel-level feasibility, licensing, cost, clinic capacity, and business strategy assessment. The completed coefficient, threshold-derived priority-set, and RF-calibration experiments demonstrate that priority coverage is stable, whereas exact selected grids remain sensitive to reasonable modeling choices; therefore, planning applications should consider a portfolio of plausible candidate locations rather than a single uniquely determined solution.

Author Contributions: Conceptualization, Shaohua Wang; methodology, Boting Xi, Shaohua Wang, Liang Long and Haojian Liang; validation, Boting Xi and Chang Liu; investigation, Siyu Zhao; resources, Shaohua Wang; data curation, Boting Xi and Siyu Zhao; writing—original draft preparation, Boting Xi; writing—review and editing, Shaohua Wang, Wenda Wang and Jingyi Zhou; supervision, Shaohua Wang; project administration, Shaohua Wang; funding acquisition, Shaohua Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant from the State Key Laboratory of Geographic Information Science and Technology, grant number 2024KFJJ009, and funded by China Southern Power Grid Company Limited, grant number ZBKJXM20240174; the National Key R&D Program of China, grant number 2023YFF0805904; the Talent Introduction Program Youth Project of the Chinese Academy of Sciences, grant numbers E43302020D, E2Z105010F; the National Natural Science Foundation of China, grant number 42471495; the Deployment Program of the Aerospace Information Research Institute, Chinese Academy of Sciences (AIRCAS), grant number E4Z202021F; and the Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, grant number E4C1020301.

Data Availability Statement: The analysis code, model scripts, and processed CSV outputs supporting the reported sensitivity and robustness analyses are publicly available through the authors'

GitHub repository: <https://github.com/HIGISX/Veterinary-Healthcare-Facilities-Location> (accessed on 25 August 2026). Raw AMap POI records are subject to the AMap Web Service API terms and therefore cannot be redistributed as unrestricted raw data. The repository provides the computational materials and processed outputs supporting the reported analyses, while the principal model parameters, random seeds, software versions, validation metrics, scenario results, and sensitivity-analysis results are also reported in the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SWCLM	Stage-Weighted Covering Location Model
MCLP	Maximal Covering Location Problem
MILP	Mixed-Integer Linear Programming
E2SFCA	Enhanced Two-Step Floating Catchment Area
RF	Random Forest
POI	Point of interest
OSM	OpenStreetMap
GIS	Geographic Information System

References

1. Neal, S.M.; Greenberg, M.J. Putting access to veterinary care on the map: A veterinary care accessibility index. *Front. Vet. Sci.* **2022**, *9*, 2022. [\[CrossRef\]](#)
2. Ng, K.Y.; Ho, C.L.; Koh, K. Spatial-temporal accessibility and inequality of veterinary service in Hong Kong: A geographic information system-based study. *Front. Vet. Sci.* **2022**, *9*, 2022. [\[CrossRef\]](#)
3. Janke, N.; Coe, J.B.; Bernardo, T.M.; Dewey, C.E.; Stone, E.A. Pet owners' and veterinarians' perceptions of information exchange and clinical decision-making in companion animal practice. *PLoS ONE* **2021**, *16*, e0245632. [\[CrossRef\]](#)
4. Applebaum, J.W.; Ellison, C.; Struckmeyer, L.; Zsembik, B.A.; McDonald, S.E. The impact of pets on everyday life for older adults during the COVID-19 pandemic. *Front. Public Health* **2021**, *9*, 2021. [\[CrossRef\]](#)
5. Holland, K.E.; Owczarczak-Garstecka, S.C.; Anderson, K.L.; Casey, R.A.; Christley, R.M.; Harris, L.; McMillan, K.M.; Mead, R.; Murray, J.K.; Samet, L.; et al. "More attention than usual": A thematic analysis of dog ownership experiences in the UK during the first COVID-19 lockdown. *Animals* **2021**, *11*, 240. [\[CrossRef\]](#)
6. Bouma, E.M.C.; Reijgwart, M.L.; Dijkstra, A. Family member, best friend, child or 'just' a pet, owners' relationship perceptions and consequences for their cats. *Int. J. Environ. Res. Public Health* **2022**, *19*, 193. [\[CrossRef\]](#)
7. Squires, R.A.; Crawford, C.; Marcondes, M.; Whitley, N. 2024 guidelines for the vaccination of dogs and cats—Compiled by the Vaccination Guidelines Group (VGG) of the World Small Animal Veterinary Association (WSAVA). *J. Small Anim. Pract.* **2024**, *65*, 277–316. [\[CrossRef\]](#)
8. Montoya, M.; Morrison, J.A.; Arrignon, F.; Spofford, N.; Charles, H.; Hours, M.-A.; Biourge, V. Life expectancy tables for dogs and cats derived from clinical data. *Front. Vet. Sci.* **2023**, *10*, 2023. [\[CrossRef\]](#)
9. Zhou, C. Exploring future GIS visions in the era of the scientific and technological revolution. *Inf. Geogr.* **2025**, *1*, 100007. [\[CrossRef\]](#)
10. Gul, M.; Guneri, A.F. Hospital location selection: A systematic literature review on methodologies and applications. *Math. Probl. Eng.* **2021**, 2021, 6682958. [\[CrossRef\]](#)
11. Rahman, M.M.; Szabó, G. Sustainable urban land-use optimization using GIS-Based Multicriteria Decision-Making (GIS-MCDM). *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 313. [\[CrossRef\]](#)
12. Toregas, C.; Swain, R.; ReVelle, C.; Bergman, L. The location of emergency service facilities. *Oper. Res.* **1971**, *19*, 1363–1373. [\[CrossRef\]](#)
13. Church, R.; ReVelle, C. The maximal covering location problem. *Pap. Reg. Sci.* **1974**, *32*, 101–118. [\[CrossRef\]](#)
14. Psyllidis, A.; Gao, S.; Hu, Y.; Kim, E.-K.; McKenzie, G.; Purves, R.; Yuan, M.; Andris, C. Points of Interest (POI): A commentary on the state of the art, challenges, and prospects for the future. *Comput. Urban Sci.* **2022**, *2*, 20. [\[CrossRef\]](#)
15. Herfort, B.; Lautenbach, S.; Porto de Albuquerque, J.; Anderson, J.; Zipf, A. A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap. *Nat. Commun.* **2023**, *14*, 3985. [\[CrossRef\]](#)
16. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)

17. Wan, S.; Chen, Y.; Xiao, Y.; Zhao, Q.; Li, M.; Wu, S. Spatial analysis and evaluation of medical resource allocation in China based on geographic big data. *BMC Health Serv. Res.* **2021**, *21*, 1084. [\[CrossRef\]](#)
18. Subal, J.; Paal, P.; Krisp, J.M. Quantifying spatial accessibility of general practitioners by applying a Modified Huff Three-Step Floating Catchment Area (MH3SFCA) method. *Int. J. Health Geogr.* **2021**, *20*, 9. [\[CrossRef\]](#)
19. Chen, L.; Chen, T.; Lan, T.; Chen, C.; Pan, J. The contributions of population distribution, healthcare resourcing, and transportation infrastructure to spatial accessibility of health care. *Inquiry* **2023**, *60*, 00469580221146041. [\[CrossRef\]](#)
20. Li, C.; Wang, J. A hierarchical two-step floating catchment area analysis for high-tier hospital accessibility in an urban agglomeration region. *J. Transp. Geogr.* **2022**, *102*, 103369. [\[CrossRef\]](#)
21. Xia, Y.; Chen, H.; Zuo, C.; Zhang, N. The impact of traffic on equality of urban healthcare service accessibility: A case study in Wuhan, China. *Sustain. Cities Soc.* **2022**, *86*, 104130. [\[CrossRef\]](#)
22. Mohammadi, A.; Mollalo, A.; Bergquist, R.; Kiani, B. Measuring COVID-19 vaccination coverage: An enhanced age-adjusted two-step floating catchment area model. *Infect. Dis. Poverty* **2021**, *10*, 118. [\[CrossRef\]](#)
23. Liu, L.; Lyu, H.; Zhao, Y.; Zhou, D. An improved Two-Step Floating Catchment Area (2SFCA) method for measuring spatial accessibility to elderly care facilities in Xi'an, China. *Int. J. Environ. Res. Public Health* **2022**, *19*, 11465. [\[CrossRef\]](#)
24. Tripathi, A.K.; Agrawal, S.; Gupta, R.D. Comparison of GIS-based AHP and fuzzy AHP methods for hospital site selection: A case study for Prayagraj City, India. *GeoJournal* **2022**, *87*, 3507–3528. [\[CrossRef\]](#)
25. Zhu, X.; Meng, X.; Zhang, M. Application of multiple criteria decision making methods in construction: A systematic literature review. *J. Civ. Eng.* **2021**, *27*, 372–403. [\[CrossRef\]](#)
26. Mortaheb, R.; Jankowski, P. Smart city re-imagined: City planning and GeoAI in the age of big data. *J. Urban Manag.* **2023**, *12*, 4–15. [\[CrossRef\]](#)
27. Bill, R.; Blankenbach, J.; Breunig, M.; Haunert, J.-H.; Heipke, C.; Herle, S.; Maas, H.-G.; Mayer, H.; Meng, L.; Rottensteiner, F.; et al. Geospatial information research: State of the art, case studies and future perspectives. *PFG* **2022**, *90*, 349–389. [\[CrossRef\]](#)
28. Cobos-Mora, S.L.; Guamán-Aucapiña, J.; Zúñiga-Ruiz, J. Suitable site selection for transfer stations in a solid waste management system using analytical hierarchy process as a multi-criteria decision analysis: A case study in Azuay-Ecuador. *Environ. Dev. Sustain.* **2023**, *25*, 1944–1977. [\[CrossRef\]](#)
29. Chen, B.; Tu, Y.; Song, Y.; Theobald, D.M.; Zhang, T.; Ren, Z.; Li, X.; Yang, J.; Wang, J.; Wang, X.; et al. Mapping essential urban land use categories with open big data: Results for five metropolitan areas in the United States of America. *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 203–218. [\[CrossRef\]](#)
30. Wu, J.; Lu, Y.; Gao, H.; Wang, M. Cultivating historical heritage area vitality using urban morphology approach based on big data and machine learning. *Comput. Environ. Urban* **2022**, *91*, 101716. [\[CrossRef\]](#)
31. Vale, D.; Lopes, A.S. Accessibility inequality across Europe: A comparison of 15-min pedestrian accessibility in cities with 100,000 or more inhabitants. *npj Urban Sustain.* **2023**, *3*, 55. [\[CrossRef\]](#)
32. Bruno, M.; Monteiro Melo, H.P.; Campanelli, B.; Loreto, V. A universal framework for inclusive 15-min cities. *Nat. Cities* **2024**, *1*, 633–641. [\[CrossRef\]](#)
33. Tariverdi, M.; Nunez-del-Prado, M.; Leonova, N.; Rentschler, J. Measuring accessibility to public services and infrastructure criticality for disasters risk management. *Sci. Rep.* **2023**, *13*, 1569. [\[CrossRef\]](#)
34. Xu, D.; Wang, S.; Liang, H.; Zhu, H.; Zhou, L.; Zhao, Z.; Zhou, J. Optimizing waste bin allocation in the Mount Yun scenic area: A data-driven approach based on the stage-weighted covering location model. *Int. J. Digit. Earth* **2026**, *19*, 2677947. [\[CrossRef\]](#)
35. Ghorbanzadeh, M.; Kim, K.; Erman Ozguven, E.; Horner, M.W. Spatial accessibility assessment of COVID-19 patients to health-care facilities: A case study of Florida. *Travel. Behav. Soc.* **2021**, *24*, 95–101. [\[CrossRef\]](#)
36. Cao, W.-R.; Shakya, P.; Karmacharya, B.; Xu, D.R.; Hao, Y.-T.; Lai, Y.-S. Equity of geographical access to public health facilities in Nepal. *BMJ Glob. Health* **2021**, *6*, e006786. [\[CrossRef\]](#)
37. McCrum, M.L.; Wan, N.; Han, J.; Lizotte, S.L.; Horns, J.J. Disparities in spatial access to emergency surgical services in the US. *JAMA Health Forum* **2022**, *3*, e223633. [\[CrossRef\]](#)
38. Wang, X.; Seyler, B.C.; Han, W.; Pan, J. An integrated analysis of spatial access to the three-tier healthcare delivery system in China: A case study of Hainan Island. *Int. J. Equity Health* **2021**, *20*, 60. [\[CrossRef\]](#)
39. Petricola, S.; Reinmuth, M.; Lautenbach, S.; Hatfield, C.; Zipf, A. Assessing road criticality and loss of healthcare accessibility during floods: The case of Cyclone Idai, Mozambique 2019. *Int. J. Health Geogr.* **2022**, *21*, 14. [\[CrossRef\]](#)
40. Giuliani, G.; Petri, E.; Interwies, E.; Vysna, V.; Guigoz, Y.; Ray, N.; Dickie, I. Modelling accessibility to urban green areas using open earth observations data: A novel approach to support the urban SDG in four European cities. *Remote Sens.* **2021**, *13*, 422. [\[CrossRef\]](#)
41. Moreno, C.; Allam, Z.; Chabaud, D.; Gall, C.; Pratlong, F. Introducing the “15-Minute City”: Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities* **2021**, *4*, 93–111. [\[CrossRef\]](#)
42. Abdelfattah, L.; Deponte, D.; Fossa, G. The 15-min city: Interpreting the model to bring out urban resiliencies. *Transp. Res. Procedia* **2022**, *60*, 330–337. [\[CrossRef\]](#)

43. Graells-Garrido, E.; Serra-Burriel, F.; Rowe, F.; Cucchiatti, F.M.; Reyes, P. A city of cities: Measuring how 15-min urban accessibility shapes human mobility in Barcelona. *PLoS ONE* **2021**, *16*, e0250080. [\[CrossRef\]](#)
44. Baobeid, A.; Koç, M.; Al-Ghamdi, S.G. Walkability and Its Relationships with Health, Sustainability, and Livability: Elements of Physical Environment and Evaluation Frameworks. *Front. Built Environ.* **2021**, *7*, 2021. [\[CrossRef\]](#)
45. Luo, W.; Wang, F. Measures of spatial accessibility to health care in a GIS environment: Synthesis and a case study in the Chicago region. *Environ. Plan. B-Urban* **2003**, *30*, 865–884. [\[CrossRef\]](#)
46. Luo, W.; Qi, Y. An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health Place* **2009**, *15*, 1100–1107. [\[CrossRef\]](#)
47. McGrail, M.R. Spatial accessibility of primary health care utilising the two step floating catchment area method: An assessment of recent improvements. *Int. J. Health Geogr.* **2012**, *11*, 50. [\[CrossRef\]](#)
48. Stacherl, B.; Sauzet, O. Gravity models for potential spatial healthcare access measurement: A systematic methodological review. *Int. J. Health Geogr.* **2023**, *22*, 34. [\[CrossRef\]](#)
49. Tatem, A.J. WorldPop, open data for spatial demography. *Sci. Data* **2017**, *4*, 170004. [\[CrossRef\]](#)
50. Haklay, M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ. Plan. B-Urban* **2010**, *37*, 682–703. [\[CrossRef\]](#)
51. Barrington-Leigh, C.; Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PLoS ONE* **2017**, *12*, e0180698. [\[CrossRef\]](#)
52. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Aroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [\[CrossRef\]](#)
53. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **2007**, *8*, 25. [\[CrossRef\]](#)
54. Delamater, P.L. Spatial accessibility in suboptimally configured health care systems: A modified two-step floating catchment area (M2SFCA) metric. *Health Place* **2013**, *24*, 30–43. [\[CrossRef\]](#)
55. Bauer, J.; Groneberg, D.A. Measuring Spatial Accessibility of health care providers—Introduction of a variable distance decay function within the Floating Catchment Area (FCA) method. *PLoS ONE* **2016**, *11*, e0159148. [\[CrossRef\]](#)
56. Shao, Y.; Luo, W. Supply-Demand Adjusted Two-Steps Floating Catchment Area (SDA-2SFCA) model for measuring spatial access to health care. *Soc. Sci. Med.* **2022**, *296*, 114727. [\[CrossRef\]](#)
57. Rossit, D.G.; Nesmachnow, S. Waste bins location problem: A review of recent advances in the storage stage of the Municipal Solid Waste reverse logistic chain. *J. Clean. Prod.* **2022**, *342*, 130793. [\[CrossRef\]](#)
58. Nevrlý, V.; Šomplák, R.; Smejkalová, V.; Lipovský, T.; Jadrný, J. Location of municipal waste containers: Trade-off between criteria. *J. Clean. Prod.* **2021**, *278*, 123445. [\[CrossRef\]](#)
59. Medrano-Gómez, X.D.; Ferreira, D.; Toso, E.A.V.; Ibarra-Rojas, O.J. Using the maximal covering location problem to design a sustainable recycling network. *J. Clean. Prod.* **2020**, *275*, 124020. [\[CrossRef\]](#)
60. Zhuo, Q.; Yan, W. Optimizing the number and location of household waste collection sites by multi-maximal covering location model: An empirical study in Minamata City, Kumamoto Prefecture, Japan. *J. Clean. Prod.* **2022**, *379*, 134644. [\[CrossRef\]](#)
61. Goodchild, M.F.; Guo, H.; Annoni, A.; Bian, L.; de Bie, K.; Campbell, F.; Craglia, M.; Ehlers, M.; van Genderen, J.; Jackson, D.; et al. Next-generation digital earth. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11088–11094. [\[CrossRef\]](#)
62. Guo, H.; Nativi, S.; Liang, D.; Craglia, M.; Wang, L.; Schade, S.; Corban, C.; He, G.; Pesaresi, M.; Li, J.; et al. Big Earth Data science: An information framework for a sustainable planet. *Int. J. Digit. Earth* **2020**, *13*, 743–767. [\[CrossRef\]](#)
63. Liang, H.; Wang, S.; Li, H.; Zhou, L.; Chen, H.; Zhang, X.; Chen, X. Sponet: Solve spatial optimization problem using deep reinforcement learning for urban spatial decision analysis. *Int. J. Digit. Earth* **2024**, *17*, 2299211. [\[CrossRef\]](#)
64. Hemmelmayr, V.C.; Doerner, K.F.; Hartl, R.F.; Vigo, D. Models and algorithms for the integrated planning of bin allocation and vehicle routing in solid waste management. *Transp. Sci.* **2013**, *48*, 103–120. [\[CrossRef\]](#)
65. Hemmelmayr, V.; Smilowitz, K.; de la Torre, L. A periodic location routing problem for collaborative recycling. *IIE Trans.* **2017**, *49*, 414–428. [\[CrossRef\]](#)
66. Owen, S.H.; Daskin, M.S. Strategic facility location: A review. *Eur. J. Oper. Res.* **1998**, *111*, 423–447. [\[CrossRef\]](#)
67. Wang, S.; Zhou, J.; Liang, H.; Zhou, L.; Wang, Z.; Zhao, S.; Wang, H. Optimization of urban service facility allocation for the SDGs: Current status, innovations, and prospects. *Sustain. Cities Soc.* **2026**, *142*, 107291. [\[CrossRef\]](#)
68. Wang, S.; Su, C.; Zhao, Z.; Liang, H.; Wang, Z.; Sun, Y.; Zhong, Y.; Zheng, J.; Huang, Z.; Zhou, J. HiSpot: An integrated and accessible python framework for spatial optimization in Urban Analytics. *Int. J. Digit. Earth* **2026**, *19*, 2687273. [\[CrossRef\]](#)

69. Clark, S.; Clarke, G.; James, W.; Lomax, N. Measuring the accessibility of veterinary care for companion animals in England and Wales. *Anim. Welf.* **2026**, *35*, e7. [[CrossRef](#)]
70. Bunke, L.; Harrison, S.; Angliss, G.; Hanselmann, R. Establishing a working definition for veterinary care desert. *JAVMA-J. Am. Vet. Med. A* **2024**, *262*, 1–8. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.