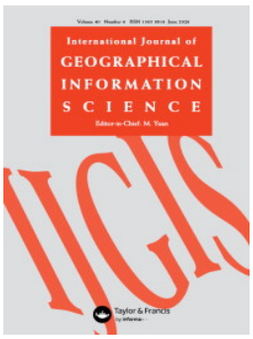


International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: www.tandfonline.com/journals/tgis20

Quantification and reduction of spatially induced uncertainty in spatial optimization via deep reinforcement learning

Gusiyuan Wang, Wangshu Mu, Changfeng Li & Yuanhui Wang

To cite this article: Gusiyuan Wang, Wangshu Mu, Changfeng Li & Yuanhui Wang (12 Jan 2026): Quantification and reduction of spatially induced uncertainty in spatial optimization via deep reinforcement learning, International Journal of Geographical Information Science, DOI: 10.1080/13658816.2025.2611984

To link to this article: https://doi.org/10.1080/13658816.2025.2611984

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2025.2611984

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Quantification and reduction of spatially induced uncertainty in spatial optimization via deep reinforcement learning

Gusiyuan Wang[a], Wangshu Mu[a], Changfeng Li[b] and Yuanhui Wang[a]

[a]Faculty of Geographical Science, Beijing Normal University, China; [b]China Academy of Urban Planning and Design, Beijing, China

ABSTRACT

Uncertainty has always been a significant yet unresolved issue in geography. Uncertainty in spatial optimization models, particularly in coverage maximization problems, stems from the limitations of binary coverage metrics for areal-based demand units. Traditional models classify units as either fully covered or not, ignoring partial coverage scenarios. This binary approach creates discrepancies between modeled and actual service demand, as even marginally covered units may count as fully serviced. The propagation of such errors is amplified by variations in the unit size and coverage thresholds, leading to suboptimal solutions in urban planning tasks such as urban greenway planning. To address these limitations, we proposed a deep reinforcement learning (DRL)-driven optimization framework that quantifies the optimality gap caused by spatially induced uncertainty and provides superior solutions. By employing a three-module structure that incorporates a graph neural network (GNN)-driven agent for data representation, a policy and a value network for decision-making, and a stochastic environment for simulating individual-level demand distributions, our model dynamically adapts to spatial uncertainty and generates robust solutions. We tested our framework in solving the maximal covering location problem for lines (MCLP-Line) for urban greenway route planning. The experimental results demonstrated that traditional mixed integer linear programming (MILP) methods yielded highly variable outcomes with different coverage metrics, whereas our proposed model consistently outperformed MILP across diverse scenarios. Our findings underscore the potential of modern AI-driven approaches in mitigating spatial uncertainty and achieving optimized solutions in geographic applications.

ARTICLE HISTORY
Received 5 June 2025
Accepted 29 December 2025

KEYWORDS
Uncertainty; coverage metric; spatial optimization; MCLP-line; deep reinforcement learning

1. Introduction

Uncertainty, defined as the discrepancy between actual and modeled spatial data, has been a persistent issue for decades in fields involving spatial data and analysis. Uncertainty can arise from various sources in spatial analysis, such as measurement

CONTACT Wangshu Mu muwangshu@bnu.edu.cn

inaccuracies, description ambiguity, and spatial aggregation. Despite the increased availability of data and techniques today, uncertainty remains prevalent at nearly every stage of spatial studies. This issue has gained prominence in disciplines such as geography (Fotheringham and Wong 1991; Dark and Bram 2007), ecology (Jelinski and Wu 1996; Comber and Harris 2022; Jiao *et al*. 2024), and epidemiology (Parenteau and Sawada 2011; Wang and Di 2020). The analysis results can be significantly affected, potentially leading to unforeseen consequences due to inappropriate findings and interpretations and, ultimately, misguide decisions and policies (Fotheringham and Wong 1991).

A critical source of uncertainty in these models stems from spatial data aggregation. For example, population data are typically aggregated into discrete areal units, such as census tracts, block groups, and blocks, due to data availability and privacy concerns. When such aggregated population data serve as demand inputs in spatial optimization frameworks, individual geolocations become approximated by their areal unit, regardless of their internal population distribution patterns. In standard coverage maximization models that treat areal units as demand distributions, binary coverage classification (fully covered or uncovered) predominates. Such implementations disregard partial coverage scenarios where spatial services reach only subsets of an areal unit's population. These uncertainties propagate through the optimization process, potentially yielding suboptimal solutions, generating plans with reduced spatial efficacy, and producing spatially infeasible solutions with rigorous constraint thresholds (Hildemann and Verstegen 2021).

The uncertainty induced by spatial aggregation is acknowledged but remains unresolved in the literature. To address this research gap, we proposed a deep reinforcement learning (DRL)-based model that focuses on assessing and reducing the impact of uncertainty caused by demand aggregation in the Maximal Covering Location Problem for Lines (MCLP-Line) (Mu and Li 2024). We tested the introduced model in an urban greenway route planning scenario in Phoenix. This research contributes to the literature in two aspects. First, the binary coverage metric with the aggregated demand representation significantly impacts the optimization results, as evidenced by varying solutions across different coverage metrics in our experiments. Using our proposed DRL framework, we quantitatively measured the optimality gap inherent in conventional models employing binary coverage assumptions. Second, the DRL model demonstrated superior performance by generating more efficient solutions and consistently outperformed traditional optimization approaches across all tested coverage metrics, highlighting its potential to address the uncertainty issue in spatial optimization problems.

The remainder of this paper is organized as follows: we start by reviewing the existing research on data uncertainty and its implications for spatial analysis. In the methods section, we introduce the technical details of our model, including problem formulation, model construction, and model training. We then conduct experiments to assess the effectiveness of our DRL model and provide a comprehensive comparison of the proposed model and the MILP models with different coverage metrics. Finally, a discussion of the model and test results, as well as the conclusion, are provided.

2. Background

Uncertainty in spatial analysis has been a persistent research focus for decades, manifesting across multiple dimensions of geospatial studies. Spatial data inherently contain measurement inaccuracies, particularly in positional data acquisition. For example, while modern global positioning system (GPS) devices typically achieve submeter accuracy ($\leq$1 m), their precision is significantly influenced by environmental factors such as signal multipath effects and atmospheric conditions (Acácio *et al*. 2022). Semantic uncertainty is another critical dimension, particularly in land cover classification. Owing to environmental variations, classification schemes inevitably differ across nations. Classes with identical names may represent different compositions of terrestrial features in different countries (Fisher 2008). This variability introduces inherent ambiguity in spatial data interpretation and cross-study comparisons. The issue of scale also receives much attention. For example, the relationship between two variables can be inconsistent at different scales of analysis in studies of public health, resulting in different inferences and conclusions (Labib *et al*. 2020).

Additionally, uncertainty can appear in almost every stage of spatial optimization, such as problem formulation, model specification, spatial attributes, location, solution and implementation (Murray 2023). Taking attributes as an example, some of the attributes can be calculated through combining two or more other attributes, which raises concerns as to whether they should be viewed as a multi-objective optimization problem involving the combined attributes. Additionally, in location decisions, the selected facilities should operate for a long duration. Thus, some of the attributes are likely to change, both during planning operations and after facilities are built. For example, the building costs could vary during construction, and the demand may change when facilities are still in service.

Spatial uncertainty in geographic analysis is manifested primarily through two well-documented problems: the Modifiable Areal Unit Problem (MAUP) and the Uncertain Geographic Context Problem (UGCoP). These issues significantly impact the reliability and validity of spatial quantitative analyses across various disciplines. The MAUP, first articulated by Openshaw (1984), arises from the inherent flexibility in defining spatial units for analysis. It comprises two interrelated components: the scale effect and the zoning effect. Aggregating smaller areal units into larger, fewer units often results in information loss, introducing uncertainty into the analysis. For example, fine-grained spatial patterns may be obscured when data are aggregated at coarser scales. Additionally, reorganizing a set of areal units into different configurations, even while maintaining the same number of units, can lead to variability in data values. This variability can produce biased conclusions, as the spatial arrangement of units directly influences analytical outcomes. A conventional example of the MAUP was provided by Openshaw (1979), who examined the relationship between Republican voters and the elderly population in Iowa. By aggregating the state's 99 counties into larger districts using different grouping methods, the correlation coefficients exhibited dramatic variation, ranging from -0.97 to $+0.99$.

The UGCoP has garnered significant attention in geography, public health, and social sciences because of its implications for understanding human behavior and experiences. This problem centers on the geographic context of individuals, which is

often oversimplified in spatial analyses. Traditional approaches rely on static residential neighborhoods as the primary contextual unit, which are typically derived from census or survey data (de Keijzer *et al*. 2020). However, individuals are mobile and frequently traverse multiple locations for work, education, recreation, and social interactions (Hamermesh *et al*. 2005). In addition, ignoring individuals' exposure to nonresidential contexts can lead to inconsistent and unreliable results. For example, studies that fail to account for daily mobility patterns may misrepresent the true geographic context influencing behavior or health outcomes (Inagami *et al*. 2007; Park and Kwan 2017).

Addressing the challenges posed by the MAUP and UGCoP requires innovative strategies. While numerous approaches have been proposed to observe the effects of these problems (Chen and Kwan 2015; Zhang *et al*. 2022; Javanmard *et al*. 2023), few comprehensive solutions have been developed. The most direct approach to mitigating the MAUP and UGCoP is to collect high-resolution spatiotemporal data that capture individual-level context by using techniques such as self-reported activity interviews (Jones and Pebley 2014), tracking devices such as handheld GPS devices and mobile phones (White *et al*. 2017; Li *et al*. 2018), and environmental sensors (Castell *et al*. 2017). These methods effectively address the MAUP by eliminating the need for aggregation and zoning, thereby reducing scale-related biases. Additionally, they enable accurate identification of micro-environments and their influence on individual behaviors and therefore also resolve the UGCoP. However, as more individuals are involved in large-scale studies, both the cost and difficulty in data management and processing significantly increase. Using individual-level data also leads to privacy concern. Without protection techniques, private information can be inferred backward through the fine-scale data collected. Developing scale-invariant methods is another way to address the uncertainties in spatial analysis. For example, by using innovative polygon-to-point distance calculation algorithms, statistical estimates based on spatially aggregated data can be consistent with those based on individual-level observations in linear regression (Mu and Tong 2020, 2022). Furthermore, the use of multiscale analysis to find the appropriate spatial scale for a particular spatial process is also feasible. Associated methods include semi-variogram analysis (Ju *et al*. 2021) and optimal parameter-based geographical detector models (He *et al*. 2023; Zhao *et al*. 2024), which have been proven to be effective in identifying the optimal spatial scale. While these methods are valuable, they are application-specific and lack universal applicability. Developing generalized solutions remains an ongoing challenge in spatial analysis.

Spatially induced uncertainty is also a recognized issue in spatial optimization, such as location analysis and modeling. Research has shown that variations in the spatial scale or unit definition can significantly affect modeling results, especially when the location is not point-based. The maximal covering location problem (MCLP), whose objective is to cover as much of the total demand as possible with a fixed number of facilities (Church and ReVelle 1974), is taken as an example. Points are commonly used to represent demands, such as population (Church 1999). However, in a large-scale optimization problem, the amount of individual data is so enormous that it is extremely difficult to solve the model in a reasonable amount of time due to computational challenges. As a result, individual data need to be aggregated and organized

according to a certain frame. Nevertheless, by varying the scale to a coarser level, an individual household may not have any significance, as multiple individual demands are combined with their neighbors and aggregated into one representation. Similar to the MAUP, different methods of aggregation may lead to errors and inaccuracies and consequently a suboptimal solution (Current and Schilling 1990). Moreover, the UGCoP is related to spatial optimization with the widespread use of specific spatial representations. The context of the representation units used in organizing and solving spatial optimization models, for example, the spatiotemporal population patterns within the units, is not well known to researchers due to abstraction and generalization during unit definition (Hu *et al*. 2024). Although distribution patterns can lead to significant differences in location problems, considering contextual properties that are not fully available in optimization models can potentially result in biased optimization results.

The binary coverage of areal units introduces significant uncertainty-induced suboptimality in spatial optimization, as varying interpretations of coverage metrics can substantially alter analytical outcomes (Murray and O'Kelly 2002). Figure 1 demonstrates this challenge, where partially covered blocks create ambiguity in determining whether they should be classified as fully covered or not. This binary approach creates uncertainty in demand estimation, as the accounted demand often diverges from actual coverage, potentially leading to substantially different optimization solutions. The practice of applying uniform coverage standards across all demand blocks, without considering their interior population distribution, further compounds this issue by introducing arbitrariness into the analysis. These discrepancies between model assumptions and real-world conditions highlight the ongoing challenge of developing coverage metrics that more accurately reflect spatial realities—a fundamental issue requiring continued attention in spatial optimization research.

Stochastic programming (Birge and Louveaux 2011) is an approach that is widely adopted to address uncertainty in optimization problems and is typically represented through uncertain parameters in mathematical formulations. When probabilistic characteristics such as the mean and variance are known, uncertainty can be modeled by corresponding probability distributions. In cases where probabilistic information is

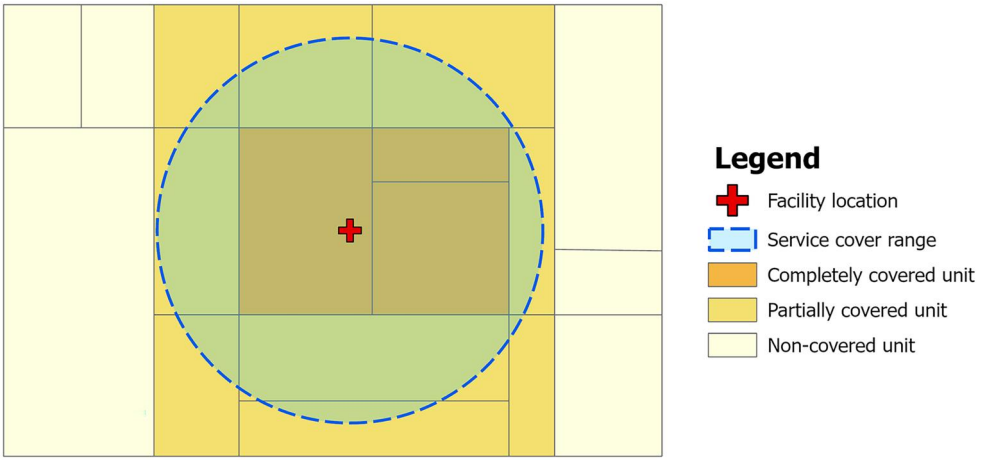


Figure 1. Complete, partial, and non-covered areal units in coverage optimization problems.

unavailable, uncertain parameters are often represented through a set of scenarios, with the objective function transformed into an expected value function across all scenarios, ensuring a total occurrence probability of 1. This scenario-based method, represented by two-stage stochastic programming, however, can be computationally intensive because of the need for numerous scenarios. In addition, applying such methods to address spatial uncertainty remains challenging, as the inherent complexity of spatial properties and relationships is difficult to encapsulate within model parameters.

Given the significant influence of uncertainty on spatial optimization problems, it is crucial to assess and mitigate its potential effects. In this research, we proposed leveraging reinforcement learning to address spatial data uncertainty in spatial optimization problems. Reinforcement learning enables an intelligent agent to learn optimal actions in a dynamic environment through trial-and-error interactions, maximizing a reward signal, which can be translated into the objective value in spatial optimization problems. The dynamic nature of the environment allows for the incorporation of uncertain input variables and the generation of samples, facilitating the simulation and evaluation of uncertainty impacts. Additionally, the agent's ability to learn through interaction with the environment enables it to produce solutions that account for uncertainty, potentially reducing its impact and yielding superior results. Reinforcement learning, combined with neural network frameworks, has already demonstrated effectiveness in solving spatial optimization problems such as the traveling salesman problem (Kool *et al*. 2018), vehicle routing problem (Zhu *et al*. 2023), and facility location problem (Liang *et al*. 2024). Furthermore, deep learning has shown promise in replacing handcrafted artificial rules (Mecheter *et al*. 2022), allowing agents to learn from uncertain environments through iterative interactions. This approach has the potential to generate solutions that outperform those derived from conventional optimization models.

This research utilized the Maximum Coverage Location Problem for Lines (MCLP-Line) model (Mu and Li 2024) as an example to demonstrate the methodology for quantifying and reducing uncertainty in spatial optimization problems. The MCLP-Line presents unique computational challenges, as it seeks to identify a contiguous, non-branching linear feature that maximizes demand coverage, where substantial overlapping service areas create modeling complexities distinct from those of conventional point-based facility location problems. While our application focused on greenway routing, where population density considerations and census data limitations introduce spatial uncertainty, the methodological contributions extend beyond urban green infrastructure planning to any linear facility location problem requiring optimization under spatial uncertainty and advancing the broader coverage location problem domain.

3. Method

In this section, we introduce a DRL model, referred to as Coverage Optimized Stochastic Reinforcement Learning for Lines (COSRL) model, designed to solve the MCLP-Line problem under demand uncertainty. In this study, uncertainty is defined as

the discrepancy between the actual demand coverage and the modeled coverage in the MILP model, arising from the use of binary coverage metrics. Such binary coverage metrics assume that coverage is either complete or absent, thereby disregarding scenarios involving partial coverage. Consequently, this methodological simplification introduces uncertainty, as the idealized coverage assumption embedded in the MILP model may not align with real-world coverage condition. To address the aforementioned issue, we developed a stochastic simulation environment that generates individual demand points. Unlike the binary coverage evaluation used in MILP models between selected lines and aggregate demand units, this environment assesses coverage based on the relationship between selected lines and the generated individual points – a level of granularity not captured by conventional MILP method. This approach mitigates the impact of partial coverage during the training process. Furthermore, we operated under the assumption of zero prior knowledge regarding the distribution of individual demands within each unit. The individual points were regenerated stochastically in each training episode, enabling the agent to learn from a diverse set of millions of potential demand distributions. As a result, the agent can produce plans that inherently account for spatial induced uncertainty.

3.1. Definition of the MCLP-line

The MCLP-Line problem is a maximal coverage location problem that focuses on linear geographic entities and aims to find an optimal single-line-shaped feature placement whose service area can cover as many demands as possible. There are existing MILP models for the MCLP-Line problem. However, the objective value of which is calculated by the sum of the demand organizing units covered by the service area, whether fully or partially. The application of such binary coverage may introduce uncertainty and result in a suboptimal result. As a result, we utilized the MCLP-Line models introduced by Mu and Li (2024) for comparison with our COSRL model. The formulation and explanation of the model can be found in Appendix A.

3.2. Markov decision process

A sequential Markov decision process (MDP) describes the interactive process between an intelligent agent and an environment, which serves as the theoretical basis for reinforcement learning. By observing the 'state' of the environment, the agent takes an 'action' and leads to a 'transition' of the environment. During this process, a 'reward' signaled by the environment is fed back to the agent, and the agent adjusts its strategy according to the received reward. In our case, two parameterized neural networks responsible for making decisions can be viewed as the agent, and the environment is the MCLP-Line problem. The geographic entities involved in the MCLP-Line problem are organized in graph form with their attributes. As shown in Figure 2, the sequential MDP contains the following components:

1. The states of the environment (s) represents the layout and corresponding attributes of the study area in the MCLP-Line problem. It is organized on the basis of

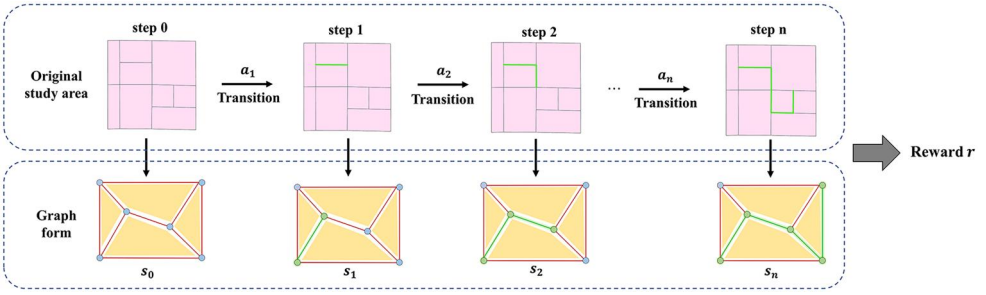


Figure 2. An example of sequential MDP.

graphs that describe the neighbor relationships of features in the study area along with other information, such as statistics of features. The original states are first sent into a graph neural network state encoder to acquire effective representations, which will be introduced later in this article.

2. Actions (a) taken by the agent indicate the choices made by the agent at every time step. The line feature segments in the MCLP-Line problem are action candidates at each timestep.
3. Rewards (r) are the feedback from the environment. The agent tries to obtain more rewards for the entire timestep. In other words, the reward encourages the agent to improve itself. In the MCLP-Line problem, we set a reward of 0 for all intermediate steps and return a reward calculated by the ratio of covered demands to total demands at the termination step.
4. Transitions (T) are made in the environment while the agent takes action. Transitions change the attributes of the graph mentioned above.

The MDP serves as the basic framework for the COSRL model. The parameterized neural networks within the agent are optimized towards covering more overall demands according to thousands of samples of states, actions, rewards, and transitions, or tuples (s, a, r, T). After a number of training iterations, the agent can generate an optimal or near-optimal solution for the MCLP-Line problem.

3.3. COSRL model

3.3.1. Overview of the model

The COSRL model comprises three modules (also see Figure 3): the presentation module, the decision module, and the evaluation module. The decision module, which constitutes the core of the COSRL model, is responsible for selecting a sequence of candidate line features to form a new simple linear feature. This process is facilitated by the collaborative operation of two DRL agents, namely, the policy network and the value network, both of which are deep neural networks. The evaluation module includes a stochastic environment capable of generating potential spatial distributions of disaggregated demands within each areal unit. In the evaluation module, we calculate the coverage of solutions provided by the decision module and offer feedback to the DRL agents. In this way, these agents can be iteratively trained to produce improved outcomes while accounting for uncertainties due to spatial aggregation.

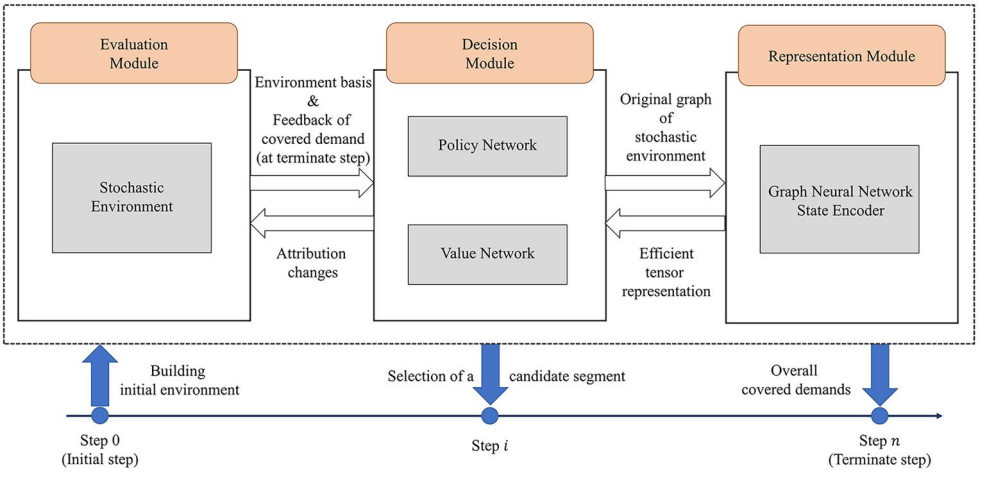


Figure 3. Overview of the COSRP model.

Additionally, the COSRL model features a graph neural network (GNN) as a presentation module that transforms networks, such as candidate road networks and solutions, into tensors, which are used as inputs by the DRL agents in the decision module. In the subsequent sections, we elaborate on the key methodologies employed in each module.

3.3.2. Evaluation module

As the key of the evaluation module, the DRL environment is the basis of reinforcement learning, which serves as a digital abstract of the problem for agents to learn. For example, in a greenway planning scenario, the environment contains candidate road segments, built greenways, census blocks, and other elements and parameters that help the agent solve the problem. To address the issue of binary coverage metric, we further designed a stochastic environment, which can provide individual-level population distributions within the areal unit on the basis of coarse population data. Figure 4 provides a comparison of the static environment and our stochastic environment. In the static environment, the coverage relationship is determined based on units. If a unit is viewed as covered under a certain binary coverage metric, all demands within the unit will be considered served. In comparison, the stochastic environment uses individual demand points in coverage measurements. Specifically, within each unit, points are generated before the training starts to indicate individual demand, whose amount is far greater than the actual population revealed by the units' attributes. The coverage relationships between the service provided by the selected linear feature and these generated points are subsequently calculated. Every time the agent finishes its episode of sampling or training, points of the same demand in each unit are chosen, and the objective can be calculated according to the coverage relationships between the points and the selected linear feature, which is measured in advance.

The introduced stochastic environment gives the DRL agent a platform to learn the uncertainty brought about by uneven individual demand distributions, which fail to

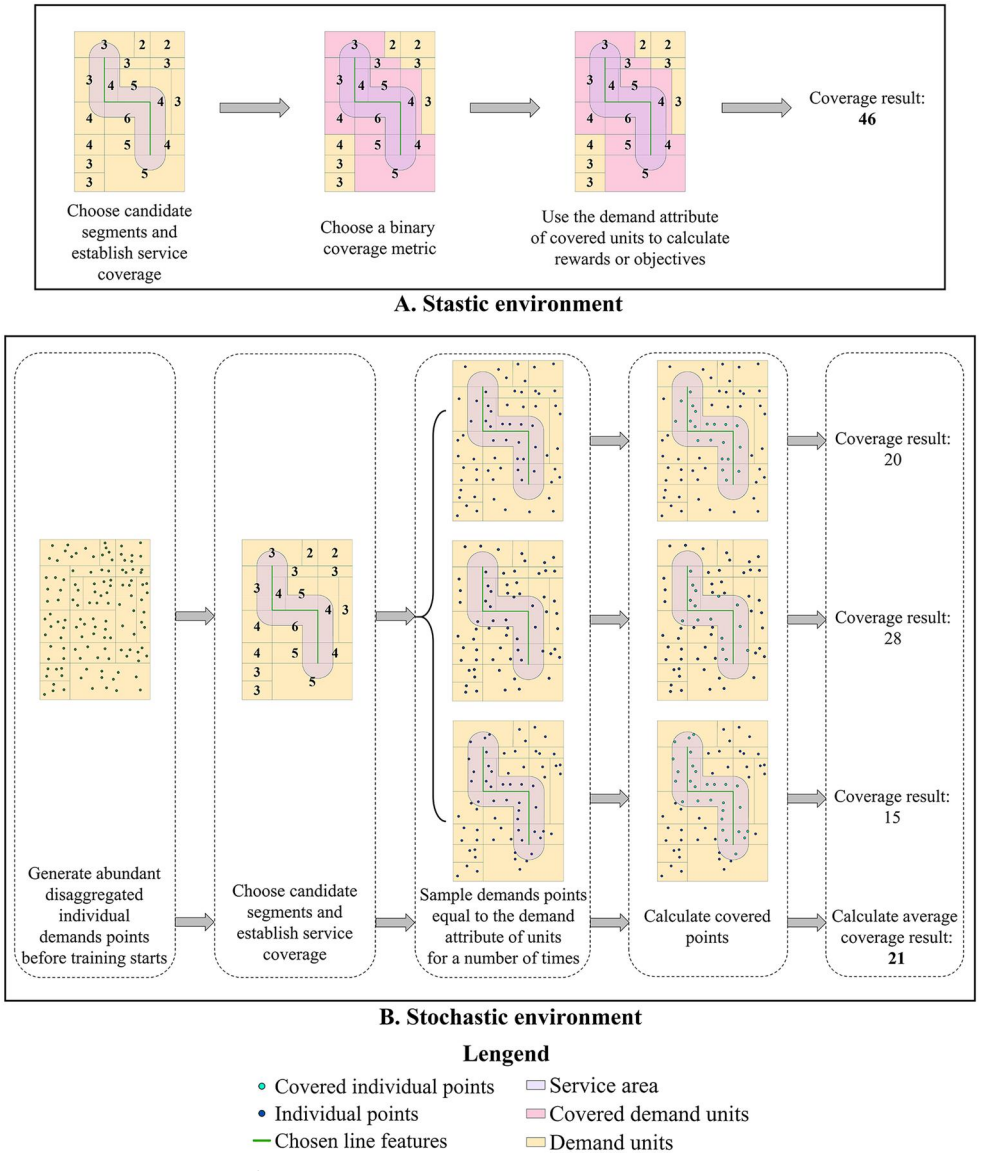


Figure 4. A comparison of the (a) static environment and (b) stochastic environment.

be exhibited on the misaligned aggregated units. The generated individual points change every time the agent restarts its interaction with the environment. Moreover, the agent can learn from countless forms of point distributions and consequently obtain a more reasonable solution than can conventional MILP models using deterministic binary coverage metrics. Additionally, the coverage measurement is conducted in advance, which can maintain efficiency during training, even in a large-scale situation.

We used individual points covered by the service area provided by selected linear feature as the reward, which is calculated at the last step as follows:

$$d_i = \min\{Dis(P_i, LF_1), Dis(P_i, LF_2), \ldots, Dis(P_i, LF_{n_{LF}})\} \quad (1)$$

$$r = \frac{\sum_{i=1}^{p} \mathbb{1}[d_i < b]}{p} \quad (2)$$

where *dis* is the Euclidean distance, d_i is the minimum distance from the generated individual point P to the linear feature LF, b is the service distance of the linear feature that can be set through our model, and p is the overall population in the study area. The reward r is calculated by the ratio of individual points covered by the service and total population. If the agent gives an invalid plan that violates the constraints in MCLP-Line, for example contains branch or self-intersection, an negative reward will be fed back to the agent.

3.3.3. Representation module

A graph neural network, as the key component of the representation module, is a specialized artificial neural network that takes graphs as input. In our case, it serves as an encoder, generating efficient representational tensors of the current stochastic environment. The environments of previous studies that applied DRL to complete tasks were typically regular and differed from the stochastic environment we aim to build. For example, a well-known case utilizing DRL is to train an agent to play the board game Go, whose environment is a 19×19 board. It can efficiently be represented by grids (Silver *et al*. 2017). Another example is playing video games such as Atari games, and the game screen as the training environment is also a grid (Mnih 2013). In contrast, the MCLP-Line problem is defined on a more complex network. The linear feature to be selected generally follows existing linear geographic entities, for example, roads. Roads and demand units can have various shapes. Additionally, they are not orthogonal in most cases, and the segments can intersect at any angle. Therefore, we need to transform complex spatial elements into simple representations.

In the MCLP-Line problem, the linear feature to be chosen is generally considered a contiguous line without branches or self-intersections, indicating that topological information is critical for the agent during training (Mu and Li 2024). Therefore, from the viewpoint of topology, we constructed a planar graph to generalize the study area, which contains demand units and candidate segments. We transformed all the geometries into elements on the graph, which include faces, edges, and nodes. The vertices and boundaries of the polygons are viewed as nodes and edges, and the polygons become faces in the graph. The attributes of the elements are shown in Table 1. Eventually, we turned different places of interest, regardless of their geometric forms, into regular graphs. The states are the attributes and topology of the current graph, and the actions are selections for the edges on the graph. During transitions, the attributes of the graph alter as the selected edge changes its type to 'selected', which also leads to subsequent changes in attributes of nodes and faces. Finally, the reward can be computed from the graph according to the attributes of faces.

Although the structure is fixed, the graph is still too complex and time-consuming for training if directly fed into the agent. Accordingly, we further developed a graph neural network encoder, which can aggregate neighborhood features and extract topology information between nodes, edges, and faces. The encoder can output representations of graphs, which is the basis for the agent to observe the state and make policy. This makes it easier for the agent to grasp the changing individual distribution

Table 1. Attributes of elements in the graph.

Element	Attribute	Data type
Node	X-coordinate	Static
	Y-coordinate	Static
	Degree centrality	Static
	Connected to selected linear feature	Dynamic
	Selected linear feature ratio	Dynamic
Edge	Length	Static
	Study area boundary	Static
	Selected segment	Dynamic
Face	Population	Static
	Covered by service	Dynamic

in the stochastic environment and benefits the learning process. The entire framework of our encoder is shown in Figure 5.

Specifically, all the features are fed into different linear layers to obtain initial embeddings, respectively:

$$n_i^0 = A_{n_i} W_n^0, e_{ij}^0 = A_{e_{ij}} W_e^0, f_k^0 = A_{f_k} W_f^0, v^0 = A_v W_v^0 \tag{3}$$

where A_{n_i}, $A_{e_{ij}}$ and A_{f_k} are the attributes of nodes, edges, and faces in the input graph, respectively; A_v represents the statistics of the environment composed of the number of edges being selected and the number of faces being covered at the current time-step; and W_n, W_e, W_f and W_v are learnable parameters. n_i is the embedding of node i, e_{ij} is the embedding of the egde between node i and node j, f_k is the embedding of face k. All the embeddings are one-dimension tensor, and the sizes of the embeddings are listed in Appendix A.

Since the agent chooses edges as its action, the aggregation of neighborhood information in our graph neural network state encoder focuses on edges. First, each edge aggregates its own embedding with the embeddings of the two connected nodes and propagates it through a linear layer and an activation layer separately. We take the average of two aggregated embeddings as the result embedding as follows:

$$e_{ij,n\to e}^{l+1} = \frac{\tanh\left(W_{n\to e}^{l+1}(n_i^l || e_{ij}^l)\right) + \tanh\left(W_{n\to e}^{l+1}(n_j^l || e_{ij}^l)\right)}{2} \tag{4}$$

where $W_{n\to e}$ is a linear transformation layer and $||$ stands for concatenation operation.

Similarly, each edge aggregates its own embedding with embeddings of adjacent faces and propagates it through a linear layer and an activation layer:

$$e_{ij,f\to e}^{l+1} = \tanh\left(W_{f\to e}^{l+1} \frac{1}{N_{F_{ij}}} \sum_{k\in F_{ij}} f_k^l\right) \tag{5}$$

where $W_{f\to e}$ is a linear transformation layer, F_{ij} is the set of adjacent faces for e_{ij}, and $N_{F_{ij}}$ is the number of faces in F_{ij}.

Since every edge also has its own attributes, we propagate initial edge embeddings through a linear layer and an activation layer as well:

$$e_{ij,e\to e}^{l+1} = \tanh\left(W_{e\to e}^{l+1} e_{ij}^l\right) \tag{6}$$

where $W_{e\to e}$ is a linear transformation layer.

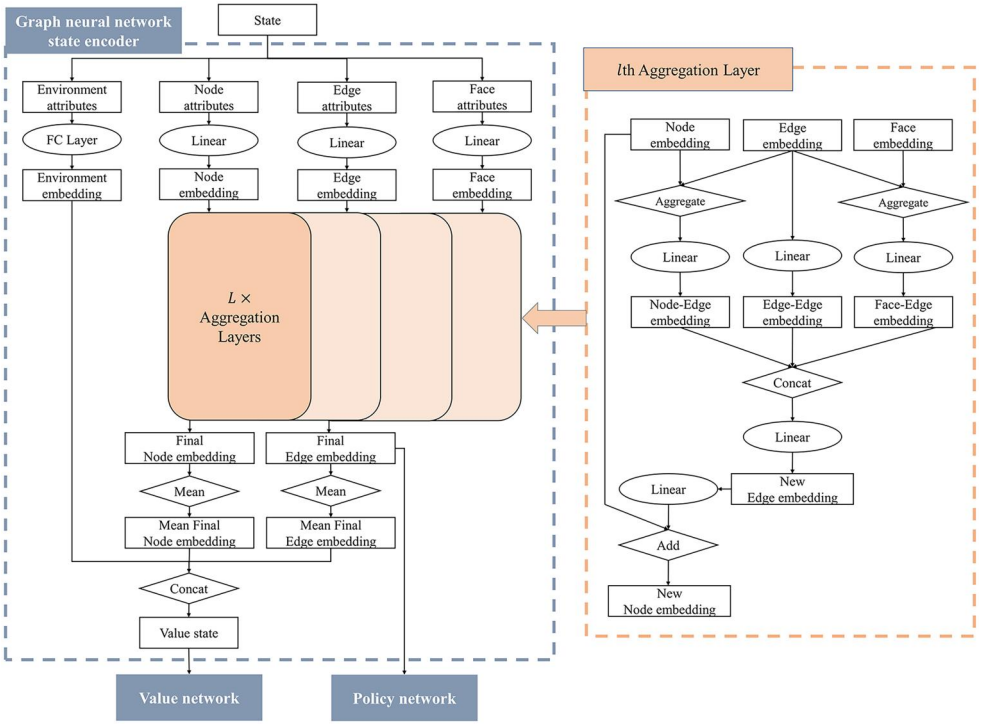


Figure 5. Overall framework of graph neural network encoder.

The embedding of edges which serves as the basis of policy-making for the agent is computed as follows:

$$e_{ij}^{l+1} = \tanh\left(W_e^{l+1}(e_{ij,n\to e}^{l+1} \,||e_{ij,e\to e}^{l+1} \,||e_{ij,f\to e}^{l+1})\right) \tag{7}$$

where W_e^{l+1} is a linear transformation layer.

Considering that nodes are involved when computing the contiguity of selected linear features, we update the node embeddings with the embedding of connected edges and add it to the node embeddings:

$$n_i^{l+1} = W_{e\to n}^{l+1} \frac{1}{N_{E_{n_i}}} \sum_{j\in E_{n_i}} e_{ij}^{l+1} + n_i^l \tag{8}$$

where $W_{e\to n}^{l+1}$ is a linear transformation layer, E_{n_i} is the set of connected edges for n_i, and $N_{E_{n_i}}$ is the number of edges in E_{n_i}.

The process above makes up a set of graph convolutional layers. After L sets of layers, we obtain e^L and n^L, final embeddings of edges and nodes, as the representations. Then, we adopt mean pooling on node and edge embeddings as follows:

$$\overline{e} = \frac{1}{M}\sum^{M} e_{ij}^L \tag{9}$$

$$\overline{n} = \frac{1}{N}\sum^{N} n_j^L \tag{10}$$

Eventually, we use a concatenation of edge embeddings, node embeddings, and embeddings of the statistics of the environment to represent the current graph completely:

$$g = v_0||\overline{e}||\overline{n} \tag{11}$$

where g is the graph embedding for the current state. Since g is formed by concatenating three embeddings, it is also a one-dimension tensor whose size equals the sum of the sizes of these three embeddings. Besides, to validate the role of representation module, we conducted an ablation experiment, which can be found in Appendix B.

3.3.4. Decision module

The decision module is composed of two neural networks, the policy network and the value network. The policy network guides the agent to take actions according to the state. The input is edge embeddings computed from the graph neural network state encoder, which generate the probability distribution for choosing edges and samples from this distribution to take action at each step. Specifically, as shown in Figure 6(a), the edges are scored through a feed-forward multi-layer perceptron (MLP) as follows:

$$q_{ij} = MLP(e_{ij}^L) \tag{12}$$

Then, we transform it into probability:

$$Prob_{ij} = Softmax(q_{ij}) \tag{13}$$

The action taken by the agent, which is edge selection, is sampled from the computed probability distribution by the policy network. Moreover, for invalid actions such as duplication or noncontiguous segments, we created an action mask to avoid these unreasonable actions, which can reduce the size of the action space and improve the efficiency of DRL algorithms. Specifically, the scored q_{ij} for invalid actions will be multiplied with a large constant negative value, making sure that the corresponding probability output by softmax being 0. Thus, only feasible actions will be allowed to be sampled.

In addition, the value network can evaluate the graph to predict the performance of the policy made by the agent and help the agent act more efficiently. As shown in Figure 6(b), it takes the complete graph embedding to obtain the current state:

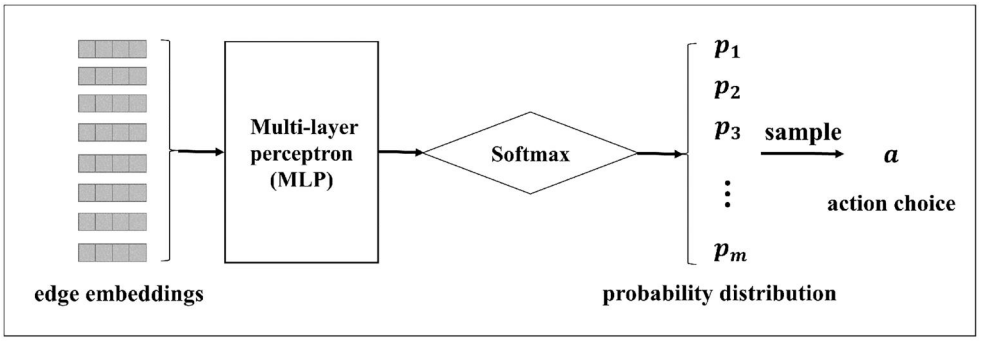
$$v_g = MLP(g) \tag{14}$$

where v_g is the estimated value of the current plan.

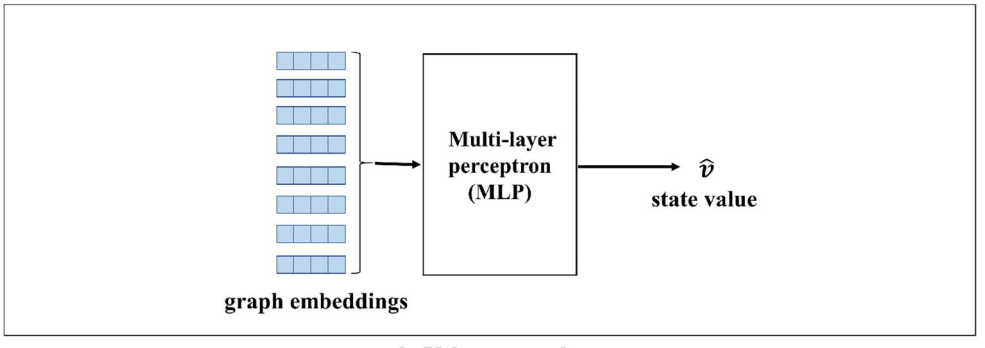
The policy network and the value network together make up the intelligent agent, and both of them will update parameters during model training.

3.3.5. Model training

Model training is composed of two processes: the sampling process and the updating process. During the sampling process, the agent ceaselessly generates plans without updating the parameters of the policy and value networks to provide abundant data for the updating process, which is generally considered helpful to intensify training (Smith *et al*. 2018). In addition, updating parameters with massive sample data can exploit the parallel computing capability of the GPU. The configurations of training parameters can be found in Appendix C.



a. Policy network



b. Value network

Figure 6. Workflow of (a) the policy network and (b) the value network.

After samples of thousands of steps in the environment are collected, we utilized proximal policy optimization to update the policy and value networks (Schulman *et al.* 2017). The loss function consists of policy loss, entropy loss, and value loss. For policy loss, it is a clipped surrogate objective function that can restrict policy updates in a safe range, which is calculated as follows:

$$r_t(\theta) = \frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta_{old}}(a_t \mid s_t)} \tag{15}$$

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t) \tag{16}$$

$$\hat{A}_t = \delta_t + (\gamma\lambda)\delta_{t+1} + (\gamma\lambda)^2\delta_{t+2} + \cdots + (\gamma\lambda)^{T-t+1}\delta_{T-1} \tag{17}$$

$$L_{policy} = \min(r_t(\theta)\hat{A}_t, clip(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t) \tag{18}$$

where θ is the parameters of our model (or the policy of the agent), $r_t(\theta)$ is the probability ratio between the new policy and old policy at time step t, r_t is the reward value at time step t which returns a zero if t is a intermediate time step, δ_t is TD error at time step t, $V(s)$ is the state value for state s, γ is the discount factor between 0 and 1, $\hat{A}_t$ is the advantage estimation at time step t, λ is a weighted parameter between 0 and 1, T is the length of trajectory, L_{policy} is the policy loss, ε is a hyperparameter, and $clip(r_t(\theta), 1-\varepsilon, 1+\varepsilon)$ restricts $r_t(\theta)$ in the range of $[1-\varepsilon, 1+\varepsilon]$. γ, λ and ε are adjustable hyperparameters in our model.

Entropy loss can control the balance between exploitation and exploration or the frequency with which the agent explores a new choice rather than choosing

the current known best option. It is calculated as follows:

$$L_{entropy} = Entropy[Prob(a_1), Prob(a_2), \ldots, Prob(a_{n_a})] \tag{19}$$

where $L_{entropy}$ is the entropy loss, $Prob(a)$ is the probability of choosing action a, which is computed in the policy network, and where n_a is the total number of actions.

We also utilize clipping to restrict the variation range for value loss, which is calculated as follows:

$$V_{targ} = \hat{A}_\theta + V_\theta \tag{20}$$

$$L_{value} = \max((V_\theta - V_{targ})^2, (clip(V_\theta, V_{\theta'} - \varepsilon, V_{\theta'} + \epsilon) - V_{targ})^2) \tag{21}$$

where L_{value} is the value loss, V_θ is the current estimated value by the value network, which is clipped around the previous value estimation $V_{\theta'}$, and ϵ is an adjustable hyperparameter.

The final loss function is a weighted sum of the three terms mentioned above:

$$L = L_{policy} + \alpha L_{entropy} + \beta L_{value} \tag{22}$$

where α and β are both hyperparameters in our model.

4. Case study

This section presents comprehensive experimental comparisons between the COSRL model and traditional MILP approaches using various binary coverage metrics across multiple urban greenway planning scenarios in Phoenix. Specifically, we evaluated three key parameters in the MCLP-Line: (1) varying greenway service cover distances, (2) different length constraints for greenway development, and (3) the availability of population distribution data at the finer census unit level. The COSRL model was implemented in Python 3.9.17 and run on a virtual machine featuring one NVIDIA RTX 4090 GPU, a 32-core vCPU at 3 GHz, and 256 GB of memory. The MILP models were solved using the IBM ILOG CPLEX 22.1.0. The solutions were visualized and analyzed in ArcGIS Pro 3.4.

4.1. Test settings

Phoenix, Arizona, the capital and most populous city in the U.S. state of Arizona, spans a total area of 1,341 km^2 and has a population of 1.65 million as of 2023. Known for its extreme heat, Phoenix experiences long, scorching summers, with 111 days annually averaging temperatures above 38 °C, primarily from May to September, with temperatures reaching 43 °C on an average of 21 days per year (du Bray *et al*. 2023). Rising temperatures, driven by climate change and the urban heat-island effect (Hsu *et al*. 2021), continue to exacerbate these conditions. Additionally, Phoenix is situated in one of the world's sunniest regions, receiving 3,872 hours of sunshine annually, comparable to the Sahara Desert (Bulk 2016).

The extreme heat and abundant sunshine in Phoenix pose significant challenges to residents (Jones *et al*. 2023), prompting the city to invest in green infrastructure to mitigate the urban heat-island effect. Greenways play a crucial role in enhancing residents' quality of life by providing cooler, shaded pathways and improving urban connectivity.

To explore this, we focus on the densely populated area of Phoenix as our study area. As illustrated in Figure 7, we utilize the population at census tracts as the demand and define their boundaries as candidate segments for greenway construction in the MCLP-Line model. The study area encompasses 83 tracts and 215 candidate segments, with a total population of 351,469. Additionally, we incorporate finer-scale census block data to evaluate the performance of the solutions, ensuring more precise and effective greenway planning. The census block-level data are not available in the training phase of the CORSL model or the solution process of the MILP model.

We further designed four coverage metrics and applied them in MILP models, which are shown in Figure 8. Specifically, any coverage metric views any partial coverage as valid coverage. For the centroid coverage metric and 50% area coverage metric, a unit is considered served if the centroid of the unit or half of the unit is covered by the service area. In the complete coverage metric, only when a unit is completely covered by the service is it considered served. To test the performance of MILP models with binary metrics under the influence of uncertain population distributions, we input their solutions into the stochastic environment and calculated the ratio of covered generated individual demands to the total population within the study area. For each solution, we repeated the test 20 times since individual data are sampled randomly. We took the average result of all tests and compared them with each other as well as with our COSRL model.

4.2. Test of different cover distances

We tested the performance of the greenway plans given by MILP models and our COSRL model by inputting the solution to the stochastic environment and calculating the coverage ratio for generated individual data. First, we fixed the maximum length of the greenways at 8000 m and used different settings of cover distance. The test results are shown in Table 2, and the maps of the solutions are illustrated from Figures 9 to 12. The 'd_c' column records the cover distance of the greenway service ranging from 800 m to 1500 m, which is in accordance with the average length of the unit boundaries (approximately 1100 m). The 'Model' column indicates the model we applied, including the MILP model with four different coverage metrics and our proposed COSRL model. The 'Time' column displays the solution time for each model. The 'Obj' column reports the objective value or the population covered by the solutions. For MILP models, it is the direct output of objective function. Foe COSRL model, it is the reward calculated by the stochastic environment. The 'Gap' column is the optimality gap for the MILP models, which is the percentage of the difference between a solution and the theoretical lower bound. The MILP model cannot further improve the solution when the gap is 0; however, the solution might not be optimal given the binary coverage metric. The last two columns show the overall population covered computed by our introduced stochastic environment according to census tract data and census block data. We use bold to highlight the model that yielded the best results in each experiment configuration. Because individual demand data are sampled randomly, we calculated each value 20 times and record their average as the result.

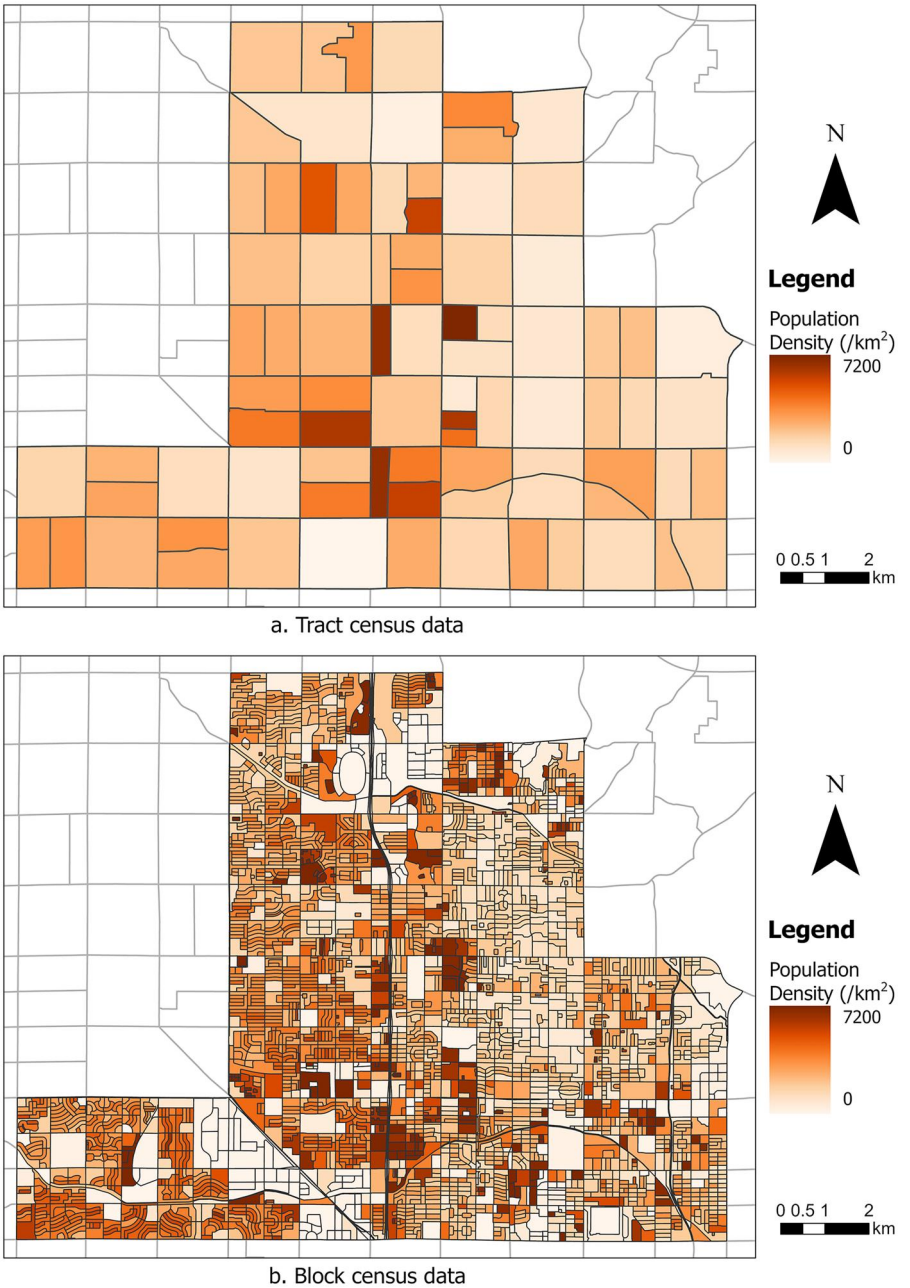


Figure 7. Study area with (a) tract-level population density and (b) block-level population density.

As shown in Table 2, the COSRL model outperformed all the MILP models when the cover distances are 800 m, 1000 m, and 1200 m. When higher-resolution data were applied, the COSRL model still yielded a superior solution. Moreover, an uneven population distribution can significantly affect MILP models with binary coverage metrics, as the objective values differed vastly from the value calculated from the stochastic environment. At the extreme, the objective value can be more than doubled or down

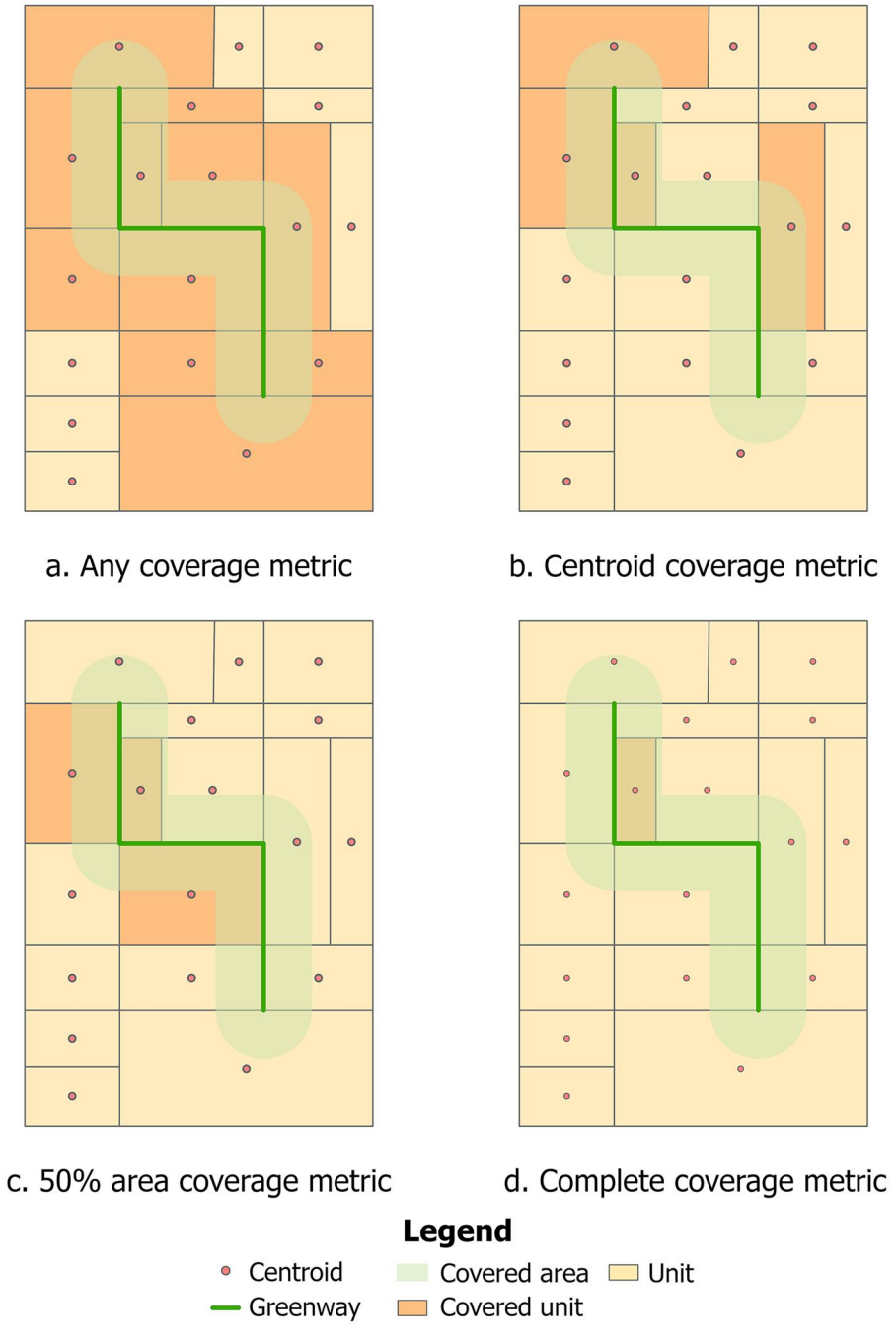


Figure 8. Binary coverage metrics in the test.

to 64% smaller compared with the value from the stochastic environment. With the increase in the cover distance, the deviation decreased in most cases. Taking the MILP model with the centroid coverage metric as an example, the deviation decreased from 23% to 15%. This could be because a greater cover distance leads to more units being

Table 2. Performance of solutions from different models with different cover distances.

d_c	Model	Time (s)	Obj.	Gap (%)	Pop (tract)	Pop (block)
800m	MILP model (any)	772	0.3253	0	0.1415	0.1461
	MILP model (centroid)	32	0.1728	0	0.1407	0.1471
	MILP model (50% area)	43	0.1475	0	0.1341	0.1315
	MILP model (complete)	7.2	0.0391	0	0.1110	0.1135
	COSRL model	0.46	0.1525	–	**0.1525 (+7.8%)**	**0.1564 (+6.3%)**
1000 m	MILP model (any)	1900	0.3757	0	0.1726	0.1796
	MILP model (centroid)	371	0.2533	0	0.1699	0.1722
	MILP model (50% area)	238	0.2533	0	0.1699	0.1722
	MILP model (complete)	3	0.1181	0	0.1742	0.1774
	COSRL model	0.42	0.1876	–	**0.1876 (+7.7%)**	**0.1924 (+7.1%)**
1200 m	MILP model (any)	1501	0.3957	≤0.01	0.2089	0.2198
	MILP model (centroid)	558	0.2852	0	0.2137	0.2219
	MILP model (50% area)	409	0.2672	0	0.2166	0.2159
	MILP model (complete)	4.7	0.1181	0	0.1949	0.1985
	COSRL model	0.45	0.2260	–	**0.2260 (+4.3%)**	**0.2289 (+3.2%)**
1500 m	Opt. model (any)	1540	0.4217	≤0.01	0.2642	0.2716
	Opt. model (centroid)	2145	0.3277	≤0.01	0.2843	0.2874
	Opt. model (50% area)	1709	0.3016	0	0.2754	0.2839
	Opt. model (complete)	30	0.1265	0	0.2467	0.2479
	COSRL model	0.45	0.2875	-	**0.2875 (+1.1%)**	**0.2874 (+0.0%)**

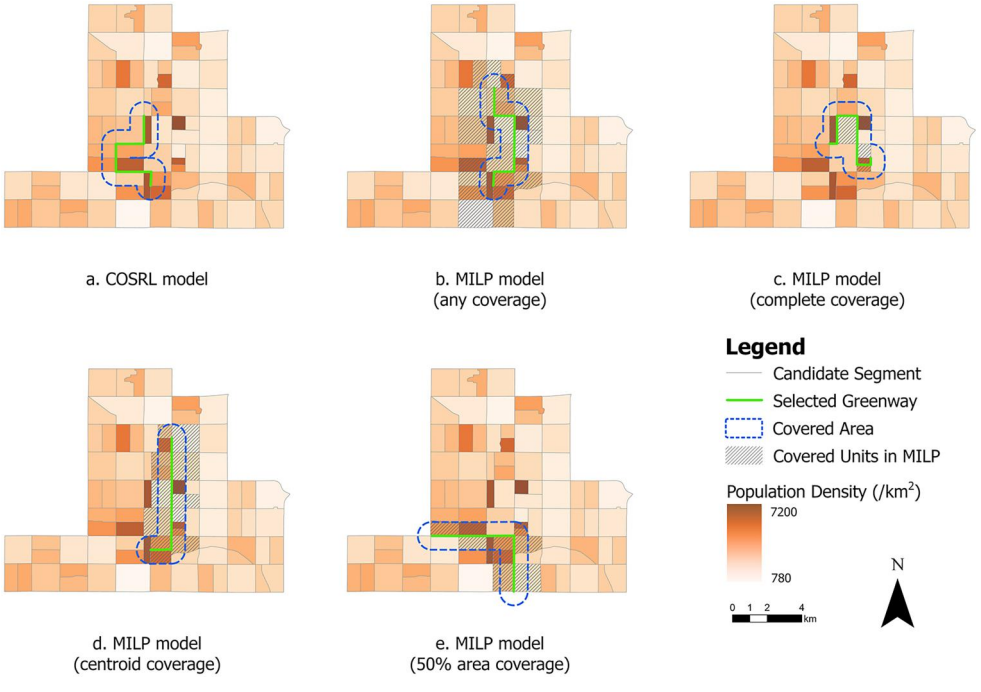


Figure 9. Solutions from different models for a maximum length of 8 km and a cover distance of 800 m.

covered mostly or completely, which decreases the degree of partial coverage. Additionally, the values calculated from census block data were also distinguished from those based on census tract data in both the COSRL model and MILP models, revealing that different data resolutions impact the results. On the other hand, the results from our COSRL model maintained outstanding performance when block data are utilized. This shows the stability of our model at a finer scale, which is not trained

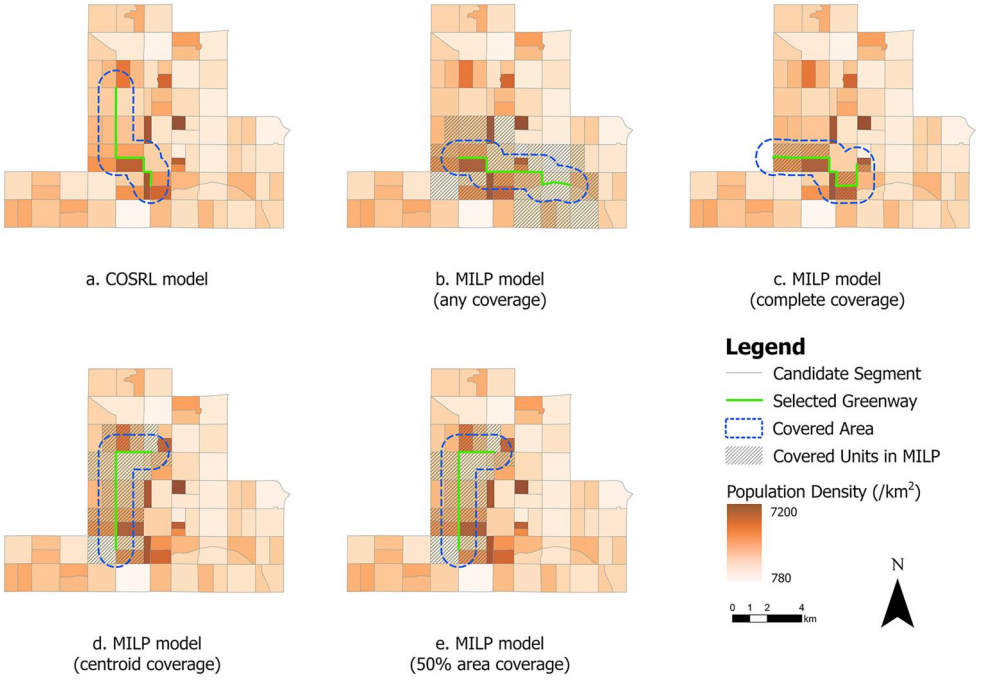


Figure 10. Solutions from different models for a maximum length of 8 km and a cover distance of 1000 m.

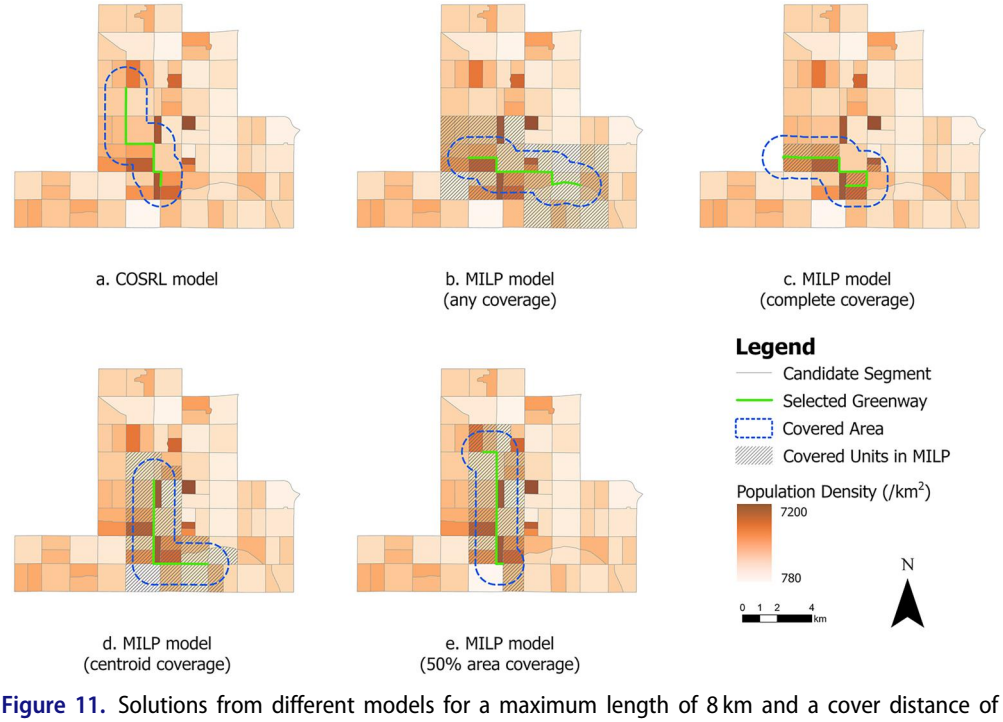


Figure 11. Solutions from different models for a maximum length of 8 km and a cover distance of 1200 m.

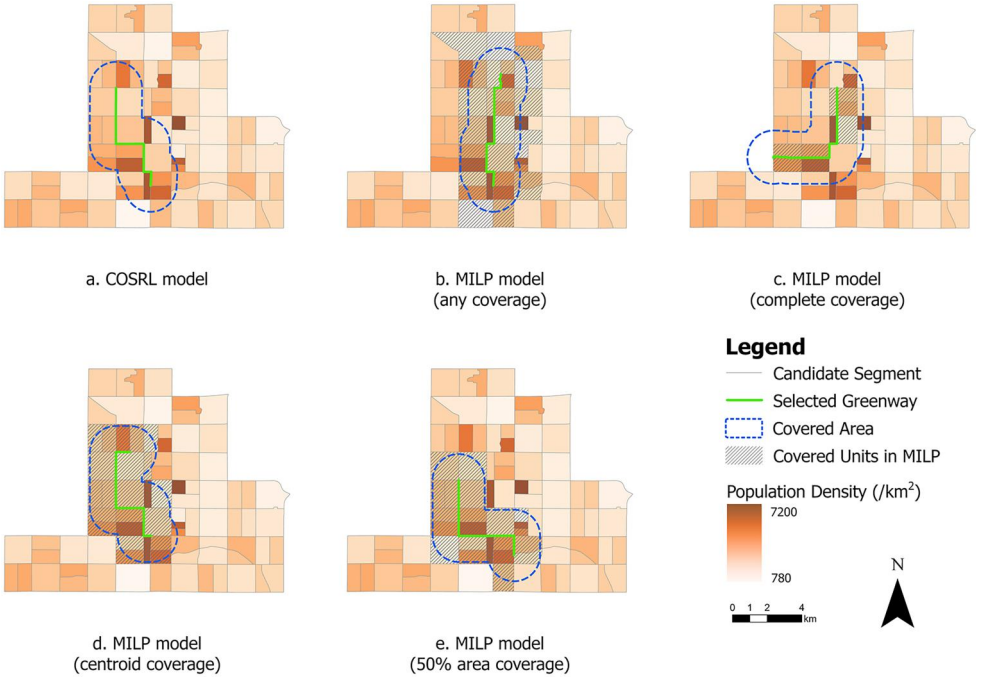


Figure 12. Solutions from different models for a maximum length of 8 km and a cover distance of 1500 m.

with during the test, demonstrating that the COSRL model can reduce the uncertainty caused by data aggregation.

4.3. Test of different maximum lengths of greenways

We then varied the maximum length of the greenways to build with a fixed cover distance of 1000 m. The test results are shown in Table 3, and the maps of the solutions are shown in Figures 13 and 14. The column 'L_{max}' records the maximum length allowed in each test. The average length of the segments in our case is approximately 1100 m, so the three settings of maximum lengths correspond to 6, 8, and 10 segments on average. Overall, our COSRL model still outperformed or yielded the same solution as MILP models do in the stochastic environment on the basis of census tract and block-level data. By comparing traditional MILP models with different binary coverage metrics, we found that similar to the previous test, the utilization of binary coverage resulted in significant variation between objective values calculated from the MILP model and our stochastic environment. The model with the complete coverage metric tended to underestimate the demand covered, whereas the others overestimated the demand covered. Consequently, although optimality gaps are almost 0, the results may still be unreliable. In addition, two binary coverage metrics, the centroid and 50% area, yielded the same solutions across all the maximum length settings, indicating considerable similarity between the two metrics. Some models with binary coverage in particular situations can generate the same plan that our model gives; however, the deviated objective value may cause a biased estimation of service-covered residents.

Table 3. Performance of solutions from different models with different maximum lengths.

L_{max}	Model	Time (s)	Obj.	Gap (%)	Pop (tract)	Pop (block)
6000m	MILP model (any)	488	0.3104	0	0.1519	0.1548
	MILP model (centroid)	114	0.2090	0	0.1358	0.1394
	MILP model (50% area)	82	0.2090	0	0.1358	0.1394
	MILP model (complete)	1.2	0.0903	0	0.1407	0.1421
	COSRL model	0.4	0.1575	–	**0.1575 (+3.7%)**	**0.1603 (+3.6%)**
8000 m	MILP model (any)	1900	0.3757	0	0.1726	0.1796
	MILP model (centroid)	371	0.2533	0	0.1699	0.1722
	MILP model (50% area)	238	0.2533	0	0.1699	0.1722
	MILP model (complete)	3	0.1181	0	0.1742	0.1774
	COSRL model	0.42	0.1876	–	**0.1876 (+7.7%)**	**0.1924 (+7.1%)**
10,000 m	MILP model (any)	5892	0.4277	$\leq$0.01	**0.2226**	**0.2262**
	MILP model (centroid)	577	0.2958	$\leq$0.01	0.2142	0.2112
	MILP model (50% area)	540	0.2856	0	0.2142	0.2112
	MILP model (complete)	5.9	0.1325	0	0.1924	0.1956
	COSRL model	0.51	0.2226	–	**0.2226 (+0.0%)**	**0.2262 (+0.0%)**

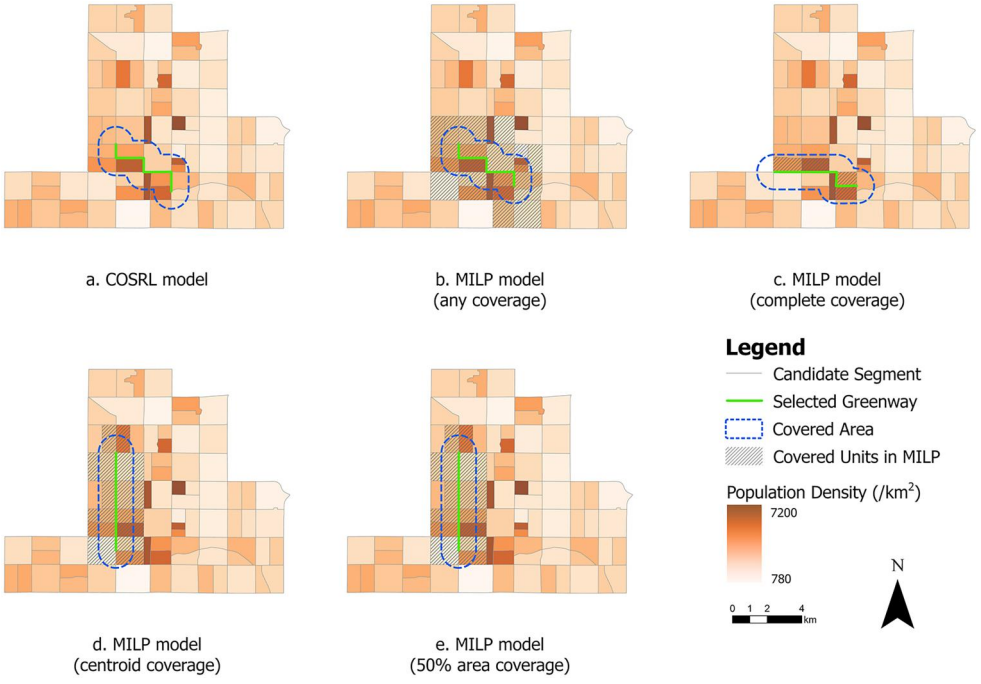


Figure 13. Solutions from different models for a maximum length of 6 km and a cover distance of 1000 m.

4.4. Test of COSRL with known demand distribution characteristics

The preceding tests assumed no prior knowledge of the demand distribution within each census tract. In practice, however, the actual population distribution can often be approximated more accurately by incorporating additional data sources, such as land use data, which may reveal non-uniform patterns. To demonstrate the performance of the COSRL model under such conditions, we simulated a non-uniform distribution with known parameters. Specifically, we used block-level data to calculate the standard deviational ellipse of the population distribution for each tract and then generated

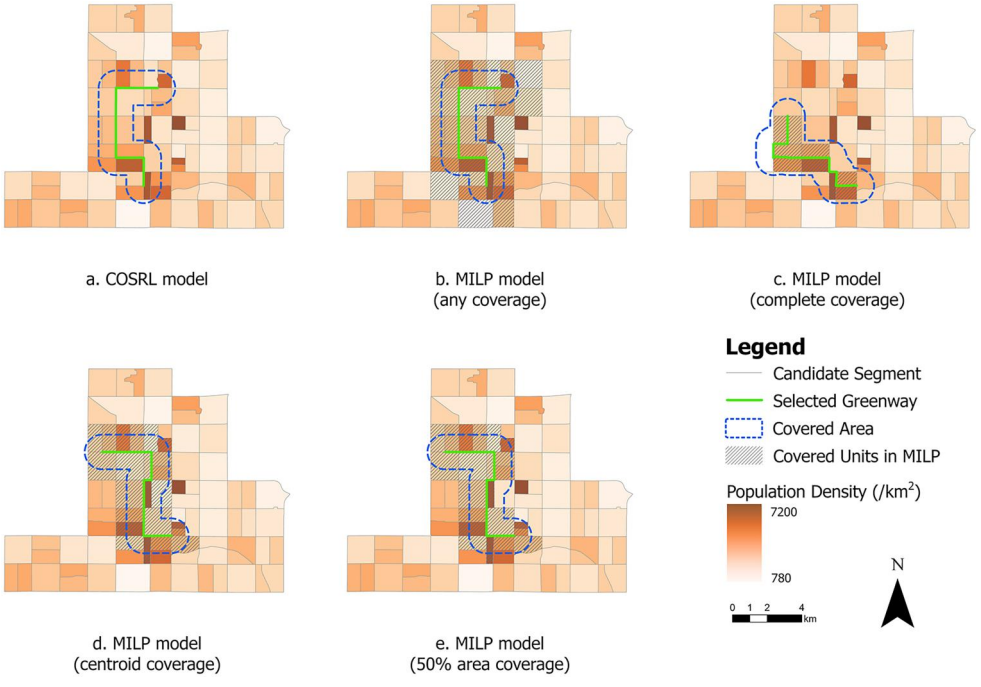


Figure 14. Solutions from different models for a maximum length of 10 km and a cover distance of 1000 m.

individual demand points accordingly. It is important to note that the agent was still trained on aggregated tract-level data; the higher-resolution block-level data were not leaked to the COSRL model and were used solely to define the parameters of the intra-tract distribution. For this experiment, the maximum greenway length was set to 8 km and the coverage distance was fixed at 1,000 m.

As presented in Table 4 and Figure 15, the results confirm that the COSRL model achieves superior outcomes when provided with more accurate population distribution data. In this test, the COSRL model effectively managed demand uncertainty and yielded solutions that surpass those of the MILP model. This advantage is notably more pronounced than in the experiment with uniformly distributed populations, indicating that our approach more effectively leverages its strengths when modeling realistic, non-uniform demand patterns. This adaptability is further evidenced by a 1% improvement in population coverage (0.1943 vs. 0.1924) validated at the block level, demonstrating that the COSRL model adjusts its solutions in response to different distribution patterns. Consequently, these findings establish that the model's performance improves significantly with access to more granular data on intra-unit demand distribution.

5. Discussions

Our experimental results demonstrate that the COSRL model outperforms MILP approaches when handling individual-level demand data, primarily because of two key advantages. First, binary coverage metrics in MILP models produce inconsistent coverage assessments, where partial coverage is either fully counted or completely ignored

Table 4. Performance of the solution from the DRL model with a finer population distribution calculated from standard deviational ellipses compared with that of MILP models.

Model	Time (s)	Obj.	Gap (%)	Pop (tract)	Pop (block)
MILP model (any)	1900	0.3757	0	0.1585	0.1799
MILP model (centroid)	371	0.2533	0	0.1623	0.1717
MILP model (50% area)	238	0.2533	0	0.1623	0.1717
MILP model (complete)	3	0.1181	0	0.1529	0.1777
COSRL model	0.48	0.1781	–	**0.1781 (+9.7%)**	**0.1943 (+8.0%)**

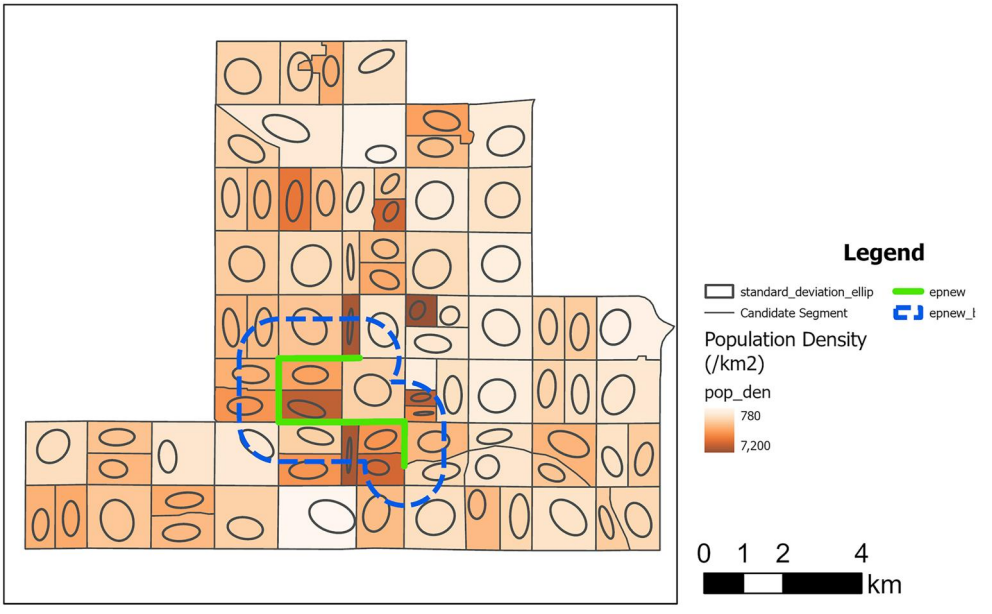


Figure 15. Solution from the COSRL model with known standard deviation ellipses.

depending on the chosen metric, leading to systematic overestimation or underestimation of true service coverage and, consequently, suboptimal solutions. Second, our COSRL model's DRL agents, trained in a stochastic environment with simulated individual-level data and coverage scenarios, learn from thousands of potential population distributions within census units – a capability beyond MILP's deterministic framework – enabling it to generate superior solutions that inherently account for spatial uncertainty while optimizing coverage objectives.

Our comparative analysis reveals important insights into model performance with various settings. When the greenway length limit and cover distance parameters are sufficiently large, both MILP and COSRL might provide the same greenway route plans, demonstrating that traditional optimization remains reliable for certain parameter configurations—although it provides misleading objective values due to binary coverage. In addition, the centroid-based and 50% area coverage metrics show remarkable consistency across all tested scenarios, with nearly identical solutions and stable performance in individual-level data, outperforming the other two coverage metrics. Moreover, incorporating detailed population distribution characteristics into COSRL would likely amplify its performance advantages over MILP by enhancing its spatial pattern adaptation capabilities.

Compared with the COSRL model, the MILP model has the advantage that less computational resources are needed, especially when the GPU is not accessible. As a result, it is necessary to choose a proper model based on the problem to solve and computing environment in practical scenarios. If partial coverage exists massively between the service area and areal units (e.g., when the cover distance is significantly lower than the average length of areal units), the COSRL model yields a more reasonable result. When precise coverage measurement is not needed and the computational device lacks a GPU, MILP models with binary coverage metrics can also be competent for solving the MCLP-Line problems.

We note that COSRL operates as a heuristic approach, inherently lacking optimality guarantees for generated solutions. To enhance solution quality, potential improvements can be applied, including architectural refinement of policy and value networks to enable more responsive agent–environment interactions through advanced reinforcement learning mechanisms, systematic hyperparameter optimization to simultaneously reduce computational overhead and improve model convergence, and augmentation of the graph neural network state encoder with spatial–relational features (e.g., candidate–demand proximity metrics) to better capture critical landscape dynamics. These enhancements strengthen the model's capacity to learn complex spatial patterns while maintaining computational tractability – a crucial balance in large-scale geographic optimization problems.

While the COSRL model effectively reduces uncertainty related to spatial representation and binary coverage limitations, its performance remains constrained by the lack of true individual-level data—a common challenge in spatial analysis owing to privacy and data availability constraints. During the generating of individual points in stochastic environment, all the points are generated uniformly, which may not be consistent with the reality. The test results of known demand distribution characteristics demonstrate that if more information of the distribution can be acquired, our COSRL model may give a completely different solution compared with uniform distribution setting and perform even better. Importantly, spatial optimization uncertainty originates from multiple interconnected sources, such as data quality, temporal variation, and model simplification, suggesting that future methodological advancements should adopt a more comprehensive approach to uncertainty quantification. This multi-faceted challenge requires continued innovation at the intersection of spatial optimization theory, computational geography, and emerging data collection techniques to develop robust solutions that address uncertainty propagation throughout the entire analytical pipeline.

The COSRL framework offers a generalized approach for mitigating the spatially induced uncertainty inherent in coverage maximization problems. It is adaptable to other line-based coverage models like the Maximum Covering Shortest Path problem with minimal reward function adjustments. Furthermore, it can be extended to optimization problems beyond coverage contexts by representing decision variables as states and actions and deriving rewards directly from objective functions. A notable advantage of this approach is its ability to implement complex constraints through procedural computer languages rather than strict mathematical formulations. However, the first and critical step for using DRL to solve spatial optimization

problems is the requisite transformation of the problem into sequential MDP, which is not always practical. A sequential MDP framework necessitates that the final solution be constructed incrementally through a series of actions. This requirement stands in direct tension with a fundamental principle of location science. As defined by Church and Murray (2008), the Laws of Location Science (LLS) posit that the sites constituting an optimal multi-site pattern are interdependent and must be selected simultaneously, not independently and sequentially. Consequently, this inherent conflict renders it particularly challenging to formulate certain problem classes – such as the p-median problem and other non-sequential point-location models – into a viable MDP. For such problems, the sequential decision-making paradigm of the COSRL model, and DRL approaches in general, may be unsuitable or require significant theoretical innovation to overcome this foundational constraint.

6. Conclusions

In this study, we introduced the COSRL model, a novel deep reinforcement learning framework, to solve the MCLP-Line problem, and overcome the limitations of MILP models that rely on binary coverage metrics for areal-based units, which often introduce uncertainties and errors. Consisting of a decision module, a representation module, and an evaluation module, the COSRL model was applied to urban greenway route planning in Phoenix across multiple scenarios. The comprehensive tests reveal distinct variations in MILP models with different binary coverage metrics, demonstrating that the use of binary coverage may misrepresent actual service demand coverage and yield unreliable results. In addition, by training an intelligent DRL agent in a stochastic environment that generates individual-level demands, the COSRL model can outperform MILP models and provide a superior solution, taking into account for uncertainty introduced by uneven population distribution within census units.

Several promising directions for future research emerge from this study. First, designing policy and value network architectures that effectively address uncertainty within DRL algorithms presents a significant challenge, requiring careful consideration of the trade-off between computational efficiency and model performance. For some large-scale problems, tailored design of graph neural networks may enhance efficiency as well. Besides, we focused on demand coverage in this research, while greenway planning needs comprehensive consideration. With data, appropriate quantization approaches and weight settings, multi-objective evaluation is able to be achieved by modifying the reward function in the COSRL model. Finally, the observed variations in objective values across MILP models employing different coverage metrics highlight the need for further investigation of the selection of appropriate metrics for the MCLP under varying conditions, building on the insights gained from our experimental results.

Acknowledgments

We would like to thank the high-performance computing support from the Center for Geodata and Analysis, Faculty of Geographical Science, Beijing Normal University.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the National Key R&D Program of China (Grant No. 2022YFC3800105) and the National Natural Science Foundation of China (Grant No. 42301476).

Data and code availability statement

The data and codes that support the findings of this study are available at https://doi.org/10.6084/m9.figshare.29261885.

Notes on contributors

Gusiyuan Wang is a master student of Geographic Information Science at Faculty of Geographical Science, Beijing Normal University. His research interests include spatial optimization and geographic artificial intelligence. His contributions include data processing, methods, experiments, writing and revisions.

Wangshu Mu is an Associate Professor in GIS at the Faculty of Geographical Science, Beijing Normal University. His research interests include spatial optimization, spatial statistics, and GIS software engineering with applications in urban studies. His contributions include conceptualization, methodology design, review and edit, and project administration.

Changfeng Li is a Senior Engineer of Urban Planning at the China Academy of Urban Planning and Design. Her research interests include spatial analysis and GeoAI methods with applications in urban planning context. Her contributions include conceptualization, review and edit.

Yuanhui Wang is a Lecturer in Human Geography at the Faculty of Geographical Science, Beijing Normal University. Her research interests focus on regional human–environment system modeling and land-use and land-cover change analysis, with an emphasis on spatial analysis and GIS-based methods. Her contributions include data curation, review and edit.

References

Acácio, M., *et al.*, 2022. Performance of GPS/GPRS tracking devices improves with increased fix interval and is not affected by animal deployment. *PloS One*, 17 (3), e0265541.

Birge, J.R., and Louveaux, F., 2011. *Introduction to stochastic programming*. New York, NY: Springer Science & Business Media.

Bulk, H., 2016. *Abstract on the Climate of Phoenix*. Available from http://www.public.asu.edu/~aunjs/ClimateofPhoenix/wxpart1.htm

Castell, N., *et al.*, 2017. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International*, 99, 293–302.

Chen, X., and Kwan, M.-P., 2015. Contextual uncertainties, human mobility, and perceived food environment: the uncertain geographic context problem in food access research. *American Journal of Public Health*, 105 (9), 1734–1737.

Church, R., and ReVelle, C., 1974. The maximal covering location problem. *Papers of the Regional Science Association*, 32 (1), 101–118.

Church, R.L., 1999. Location modelling and GIS. *Geographical Information Systems*, 1, 293–303.

Church, R.L., and Murray, A.T., 2008. Introduction. In: *Business site selection, location analysis and GIS*. Hoboken, NJ: Wiley, 1–17.

Comber, A., and Harris, P., 2022. The importance of scale and the MAUP for robust ecosystem service evaluations and landscape decisions. *Land*, 11 (3), 399.
Current, J.R., and Schilling, D.A., 1990. Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis*, 22 (2), 116–126.
Dark, S.J., and Bram, D., 2007. The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography: Earth and Environment*, 31 (5), 471–479.
de Keijzer, C., Bauwelinck, M., and Dadvand, P., 2020. Long-term exposure to residential greenspace and healthy ageing: a systematic review. *Current Environmental Health Reports*, 7 (1), 65–88.
du Bray, M.V., *et al.*, 2023. Beyond extreme: heat emergency and water insecurity for people experiencing houselessness in Phoenix, Arizona, USA during and after the heatwave of 2023. *Human Ecology*, 51 (5), 799–808.
Fisher, P., 2008. Uncertainty, semantic. In: S. Shekhar and H. Xiong, eds. *Encyclopedia of GIS*. New York, NY: Springer US, 1194–1196.
Fotheringham, A.S., and Wong, D.W.S., 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A: Economy and Space*, 23 (7), 1025–1044.
Hamermesh, D.S., Frazis, H., and Stewart, J., 2005. Data watch the American time use survey. *Journal of Economic Perspectives*, 19 (1), 221–232.
He, Q., *et al.*, 2023. Spatial stratified heterogeneity and driving mechanism of urban development level in China under different urban growth patterns with optimal parameter-based geographic detector model mining. *Computers, Environment and Urban Systems*, 105, 102023.
Hildemann, M., and Verstegen, J.A., 2021. Quantifying uncertainty in Pareto fronts arising from spatial data. *Environmental Modelling & Software*, 141, 105069.
Hsu, A., *et al.*, 2021. Disproportionate exposure to urban heat island intensity across major US cities. *Nature Communications*, 12 (1), 2721.
Hu, Q., *et al.*, 2024. Multi-scale population analysis unit construction method considering scene feature variability and long/short-term patterns in spatiotemporal population activities. *International Journal of Digital Earth*, 17 (1), 2358850.
Inagami, S., Cohen, D.A., and Finch, B.K., 2007. Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Social Science & Medicine (1982)*, 65 (8), 1779–1791.
Javanmard, R., *et al.*, 2023. The impacts of the modifiable areal unit problem (MAUP) on social equity analysis of public transit reliability. *Journal of Transport Geography*, 106, 103500.
Jelinski, D.E., and Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, 11 (3), 129–140.
Jiao, Z., *et al.*, 2024. A framework for analyzing the effects of modifiable areal unit problem on ecological security pattern. *Journal of Cleaner Production*, 458, 142549.
Jones, A., *et al.*, 2023. Climate change impacts on future residential electricity consumption and energy burden: a case study in Phoenix, Arizona. *Energy Policy*, 183, 113811.
Jones, M., and Pebley, A.R., 2014. Redefining neighborhoods using common destinations: social characteristics of activity spaces and home census tracts compared. *Demography*, 51 (3), 727–752.
Ju, H., *et al.*, 2021. Spatiotemporal patterns and modifiable areal unit problems of the landscape ecological risk in coastal areas: a case study of the Shandong Peninsula, China. *Journal of Cleaner Production*, 310, 127522.
Kool, W., Hoof, H. v., and Welling, M., 2018. Attention, learn to solve routing problems! arXiv. arXiv:1803.08475.
Labib, S.M., Lindley, S., and Huck, J.J., 2020. Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment. *Computers, Environment and Urban Systems*, 82, 101501.
Li, D., *et al.*, 2018. Moving beyond the neighborhood: daily exposure to nature and adolescents' mood. *Landscape and Urban Planning*, 173, 33–43.
Liang, H., *et al.*, 2024. Sponet: solve spatial optimization problem using deep reinforcement learning for urban spatial decision analysis. *International Journal of Digital Earth*, 17 (1), 2299211.

Mecheter, I., *et al.*, 2022. Deep learning with multiresolution handcrafted features for brain MRI segmentation. *Artificial Intelligence in Medicine*, 131, 102365.
Mnih, V., 2013. *Playing Atari with deep reinforcement learning*. arXiv. arXiv:1312.5602.
Mu, W., and Li, C., 2024. Optimization of urban greenway route using a coverage maximization model for lines. *Computers, Environment and Urban Systems*, 112, 102155.
Mu, W., and Tong, D., 2020. Distance in spatial analysis: measurement, bias, and alternatives. *Geographical Analysis*, 52 (4), 511–536.
Mu, W., and Tong, D., 2022. Computation of the distance between a polygon and a point in spatial analysis. *International Journal of Geographical Information Science*, 36 (8), 1575–1600.
Murray, A.T., 2023. Sources of uncertainty in location analysis. In: H. A. Eiselt and V. Marianov, eds. *Uncertainty in facility location problems*. Cham, Switzerland: Springer International Publishing, 3–24.
Murray, A.T., and O'Kelly, M.E., 2002. Assessing representation error in point-based coverage modeling. *Journal of Geographical Systems*, 4 (2), 171–191.
Openshaw, S., Taylor, P.J., and Wrigley, N., 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *In: Statistical applications in the spatial sciences*. London, UK: Pion, 127–144.
Openshaw, S., 1984. The modifiable areal unit problem. Norwich, UK: Geo Books.
Parenteau, M.-P., and Sawada, M.C., 2011. The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *International Journal of Health Geographics*, 10 (1), 58.
Park, Y.M., and Kwan, M.-P., 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place*, 43, 85–94.
Schulman, J., *et al.*, 2017. *Proximal policy optimization algorithms*. arXiv. arXiv:1707.06347.
Silver, D., *et al.*, 2017. Mastering the game of go without human knowledge. *Nature*, 550 (7676), 354–359.
Smith, S.L., *et al.*, 2018. *Don't decay the learning rate, increase the batch size*. arXiv. arXiv: 1711.00489.
Wang, Y., and Di, Q., 2020. Modifiable areal unit problem and environmental factors of COVID-19 outbreak. *The Science of the Total Environment*, 740, 139984.
White, M.P., *et al.*, 2017. Natural environments and subjective wellbeing: different types of exposure are associated with different aspects of wellbeing. *Health & Place*, 45, 77–84.
Zhang, S., *et al.*, 2022. Do spatiotemporal units matter for exploring the microgeographies of epidemics? *Applied Geography (Sevenoaks, England)*, 142, 102692.
Zhao, X., *et al.*, 2024. Quantitative analysis of fractional vegetation cover in southern Sichuan urban agglomeration using optimal parameter geographic detector model, China. *Ecological Indicators*, 158, 111529.
Zhu, T., *et al.*, 2023. An accelerated end-to-end method for solving routing problems. *Neural Networks: The Official Journal of the International Neural Network Society*, 164, 535–545.

Appendix A. The MCLP-Line model

The MILP models used in the tests of the urban greenway routing problem are built on the existing MCLP-Line model (Mu and Li 2024). Consider the following variables:

i, j: index of candidate road segments.
k: index of areal units.
τ: index of road junctions.
n: total number of candidate segments.
m: total number of areal units.
D: the maximum length of greenways to be built.
d_i: the length of the road i.
c_k: the population within an areal unit k, representing the demand of the areal unit k.
N_k: set of road indices whose service area can cover areal unit k.
Ω_i: set of road segment indexes that connected with the road i.

R_τ: set of road segments that connect to the road junction or ending point τ.
M: a large number.
D: the maximum length of greenways to be built.
The decision variables include:

$$x_i = \begin{cases} 1, & \text{if road segment } i \text{ is selected as a part of the urban greenway} \\ 0, & \text{otherwise} \end{cases}$$

$$y_k = \begin{cases} 1, & \text{if areal unit } k \text{ is covered by the urban greenways service area} \\ 0, & \text{otherwise} \end{cases}$$

$$s_i = \begin{cases} 1, & \text{if road } i \text{ is the starting segment of the urban greenway} \\ 0, & \text{otherwise} \end{cases}$$

$$w_i = \begin{cases} 1, & \text{if road } i \text{ is the ending segment of the urban greenway} \\ 0, & \text{otherwise} \end{cases}$$

$$f_{ij} = \begin{cases} 1, & \text{road } i \text{ and j are connected on the selected urban greenway in the road network} \\ 0, & \text{otherwise} \end{cases}$$

u_i: the order of the road segment i on the selected urban greenway, $u_i \in \mathbb{Z}^+$
The MCLP-Line model is specified as:
Maximize:

$$\sum_{k=1}^{m} c_k y_k \tag{A1}$$

Subject to:

$$\sum_{i \in N_k} x_i \geq y_k, \quad \forall k \tag{A2}$$

$$\sum_{i=1}^{n} d_i x_i \leq D \tag{A3}$$

$$s_i \leq x_i, \quad \forall i \tag{A4}$$

$$\sum_{i=1}^{n} s_i = 1 \tag{A5}$$

$$w_i \leq x_i, \quad \forall i \tag{A6}$$

$$\sum_{i=1}^{n} w_i = 1 \tag{A7}$$

$$s_i + w_i \leq 1, \quad \forall i \tag{A8}$$

$$\sum_{i=1}^{n} \sum_{j \in \Omega_i} f_{ij} = \sum_{i=1}^{n} x_i - 1 \tag{A9}$$

$$\sum_{j \in \Omega_i} f_{ij} \leq x_i, \quad \forall i \tag{A10}$$

$$\sum_{i \in \Omega_j} f_{ij} \leq x_j, \quad \forall j \tag{A11}$$

$$\sum_{j \in \Omega_i} f_{ij} - \sum_{j \in \Omega_i} f_{ji} = s_i - w_i, \quad \forall i \tag{A12}$$

$$\sum_{i \in R_\tau} x_i \leq 2, \quad \forall \tau \tag{A13}$$

$$u_i - u_j + M f_{ij} \leq M - 1, \quad \forall i, j; i \neq j \tag{A14}$$

$$1 \leq u_i \leq \sum_{i=1}^{n} x_i, u_i \in \mathbb{Z} \tag{A15}$$

$$x_i, y_k, w_i, s_i, f_{ij} \in \{0, 1\} \tag{A16}$$

Equation (1), as the objective function, is to maximize the overall population served by the selected urban greenway, calculated by the areal units covered by greenway service area. Constraints (2) ensure that areal units will be considered served if covered by a built greenway. We note that N_k, set of road indices whose service area can cover areal unit k, is determined ahead by a certain binary coverage metric. Constraints (3) specify the length limit for the selected urban greenway. Constraints (4) to (7) ensure that there are exactly one starting segment and one ending segment for each greenway. Constraints (9) to (12) applied a flow model to guarantee continuity of the greenway without any breaks. Constraints (13) ensure that the urban greenway will not self-intersect. Constraints (14) and (15) employ Miller–Tucker–Zemlin (MTZ) style subtour elimination constraints. Constraint (16) defines binary variables.

Appendix B. Ablation experiment on representation module

To further validate the role of the representation module in our model, we supplemented our analysis with an ablation experiment. We conducted two experiment, one with representation enabled and the other remove it from the COSRL model, which directly passes the attributes of the spatial entities to the agent without any encoding. In the experiment where the representation module was disabled, the agent's performance plummeted, with the final solution showing a 22% reduction in population coverage compared to the model with the representation module enabled (Figures B1 and B2). Meanwhile, the reward convergence exhibited greater instability. This experiment confirms the critical importance of the representation module in our framework. Without it, our model would be incapable of achieving the presented experimental outcomes.

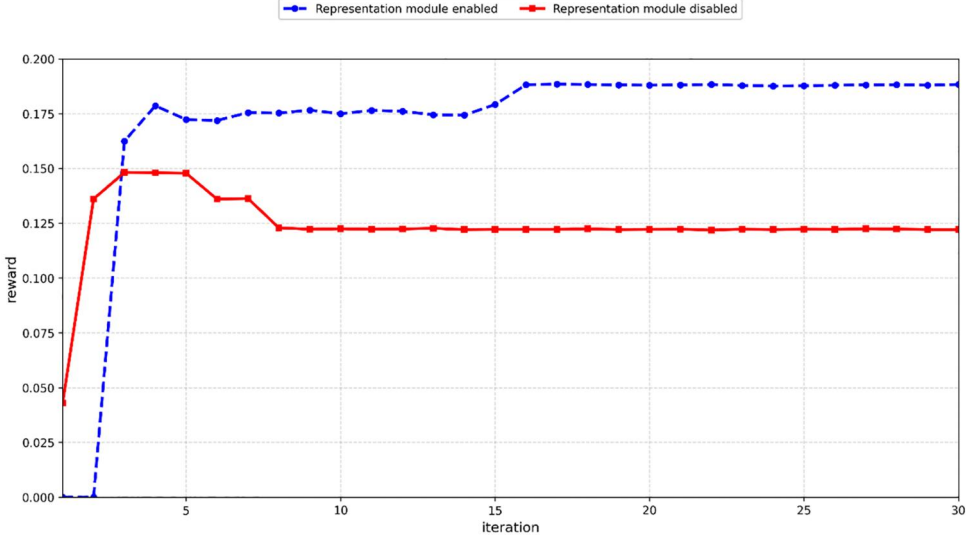


Figure B1. Reward values during training with the representation module enabled and disabled.

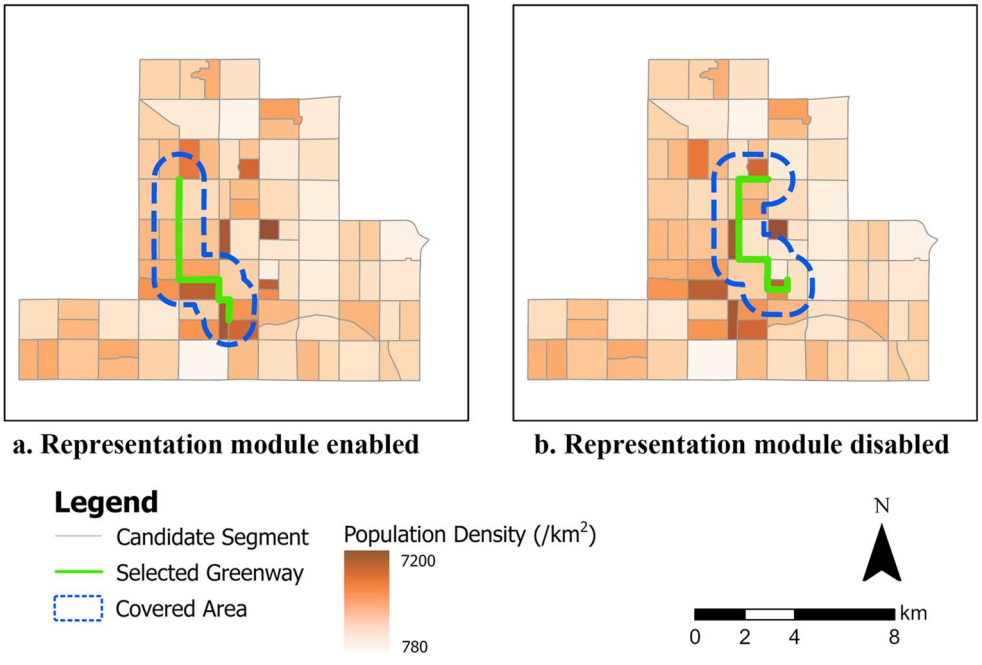


Figure B2. Solutions output from COSRL model with representation module enabled and disabled.

Appendix C. Hyper-parameters and training setting of the COSRL model

We programmed and ran the model in Python 3.9.17, and the source codes and trained model can be found at https://doi.org/10.6084/m9.figshare.29261885. We trained the COSRL model on a virtual machine featuring one NVIDIA RTX 4090 GPU, a 32-core vCPU at 3 GHz, and 256 GB of memory. All the checkpoints provided were trained for 50 iterations, which took around 18 hours. The settings of hyper-parameters and other training parameters in our model during training are listed as follows:

Table C1. Settings of hyper-parameters and training parameters.

Category	Name	Value
GNN encoder	Embeddings dimension for nodes	16
	Embeddings dimension for edges	8
	Embeddings dimension for faces	8
	Embeddings dimension for environment attributes	4
	Environment attributes MLP width	16
	GNN layer depth	1
Policy network	Policy MLP width	32
	Policy MLP depth	1
Value network	Value MLP width	32
	Value MLP depth	2
PPO	Discount factor γ	0.99
	GAE λ	0.95
	Clipping ϵ	0.2
	Entropy loss weight α	0.01
	Value loss weight β	0.5
	Gradient clipping	0.5
Training	Seed	4
	Epochs used for training	5
	Batch size	50000
	Mini batch size	1024
	Optimizer	Adam
	Adam eps	0.00005
	Learning rate	0.00035
	Times of individual demand generated before training	5
	Number of tests	50