Research Journal in Civil, Industrial and Mechanical Engineering
ISSN: 3079-2630
DOI: https://doi.org/10.61424/rjcime.v3i1.1143
Date Published Online: November 08, 2026
Journal Homepage: https://bluemarkpublishers.com/index.php/RJIEE

RJCIME
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Multi-Objective Deep Reinforcement Learning-Based Spatial Optimization of Green Infrastructure in Rapidly Densifying Cities: A Case Study of Ghorahi, Dang, Nepal

Er. Bikram Babu KC[1] ✉ and Surendra Pal Singh[2]
[1]Dr. K.N. Modi University; Lecturer, REC, Pokhara University, Nepal
[2]Professor & HOD, Department of Civil Engineering, Dr. K.N. Modi University, Newai. Rajasthan, India
Corresponding Author: Er. Bikram Babu KC, **E-mail**: bikram.babu11@gmail.com

| ABSTRACT
Urban microclimatic degradation is compounded by the rate of rapid urban expansion and rapidly increasing urban density in rapidly densifying metropolitan areas, which increase the Urban Heat Island (UHI) effect and worsen the disruption of ecological connectivity. Two-dimensional (2D) spatial planning currently existing in the planning framework has not considered the vertical aspect of urban densification and the vertical dimension also influences the sub-optimal location of Green Infrastructure (GI) and Blue-Green Infrastructure (BGI). To address the above limitations this paper presents a novel three-dimensional (3D) spatial optimization algorithm, which integrates high-resolution remote sensing data with 3D morphology and microclimate models generated by UAVs and multi-objective Deep Reinforcement Learning (DRL). The methodology is based on the formulation of GI placement as a Markov Decision Process (MDP) so that three project aspects, namely 3D microclimatic cooling benefit, pedestrian heat stress and ecological network connectivity will be optimized without being subjected to socio-economic fairness conditions, without the risk of algorithmic bias and spatial displacement. The proposed 3D-aware DRL optimization has reduced heat stress by a mean of 2.8 °C within the urban testbed area of 5 km^2 in Ghorahi Sub-Metropolitan City, Dang, Nepal from conventional 2D heuristic of 1.2 °C compared to base condition and increased the Integral Index of Connectivity (IIC) from 0.14 to 0.45 compared to base condition. The framework further translates model outputs into actions by linking with urban governance tools using the municipal building bylaw and requirement of green space coverage provided in the Local Government Operation Act (2017) in the municipalities of Nepal through aligning outputs of scenarios with municipal bylaws. This work merges high-dimensional and AI-driven analytics and socio-technical urban governance and offers a scalable blueprint for urban densification that is robust, climate-resilient and equitable indeed.

| KEYWORDS

Green Infrastructure, Deep reinforcement learning, UHI – Urban Heat Island, three-dimensional Microclimate, Spatial optimization, Algorithmic fairness, Blue-Green Infrastructure, Urban morphology and Nepal.

| ARTICLE INFORMATION

ACCEPTED: September 01, 2026 **PUBLISHED:** November 08, 2026 **DOI:** https://doi.org/10.61424/rjcime.v3i1.1143

1. Introduction

The density of urbanization is becoming the key characteristic of the 21st century. Spatial scarcity in urban areas (in particular in the rapidly "up-growing" urban regions) forces urban built environments to be squeezed and to grow upwards, thereby altering the 3D parameters of the urban canopy, radiation trapping and wind flow regimes. Urbanisation effects: The impact of the changes in structure increases the Urban Heat Island effect, the heat stress

Copyright: © 2026 The Author(s). This open access article is distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC- BY-NC-SA) License (https://creativecommons.org/licenses/by-nc-sa/4.0/), published by Bluemark Publishers.

of pedestrians, air quality and the fragmentation of the urban ecological networks. Important nature-based solutions for urban climate adaptation and provision of ecosystem services are all considered, and are referred to as Green (GI) and Blue-Green (BGI) Infrastructure, including parks, green roofs, vertical green walls, street trees and urban wetlands.

Traditional urban spatial planning, however, is mostly based on 2D spatial analysis, land-surface temperature (LST) mapping and static land-suitability overlays. It can be seen from the current situation of the research of AI-based BGI models that there are still many knowledge gaps: indeed Ignoring 3D bioclimatic factors in current machine learning models: The majority of the currently available machine learning models rely on 2D morphology and surface temperature, but not 3D volumetric morphological information and vertical shading as well as air temperature and the Universal Thermal Climate Index (UTCI) that quantifies human thermal comfort.

Appropriate model transferability and dynamic optimization: Existing optimization models are not transferable to complex urban topographies and are not able to support the evaluation of dynamic scenarios with changing climate risks.

Potential spatial injustice in the normal algorithmic optimisation processes: Algorithmic bias in the optimisation processes can lead to further spatial injustice, such as a more advantageous real estate location, that may exclude vulnerable groups even more, and result in green gentrification. Lack of socio-technical governance and translation of technical optimisation products to development-control mechanisms.

To address these challenges, in this paper a complete multi-objective DRL approach for rapid densification of an urban area is presented, and is tested empirically in Ghorahi Sub-Metropolitan City of Dang district in Nepal. The aim is to describe a full solution including high dimensional remote sensing data and integration of multi-criteria policy implementation to maximize the cooling, biodiversity connectivity and social equity at the microclimatic scale.

2. Literature Review and Conceptual Framework

This research is based on a theoretical framework that encompasses landscape ecology, low carbon evaluation in urban planning using machine learning and socio-ecological-technical systems (SETS). Spatial connectivity (Ahern, 2007) and multi-functionality and pattern-process relationship across spatial scales (Lafortezza et al., 2013) were the two main principles that underpinned early planning scholarships in the study of GI. With the growing extent of urban planning, there have been a growing number of machine learning applications to climate decision support models (CDSM) that process spatial data (Milojevic-Dupont & Creutzig, 2019).

For the vertically complex urban context, the performance of the GIs is closely related to the 3D building morphology, sky view factors (SVF), aspect ratio and dynamics of the solar envelope. Over the past few years, there have been a lot of developments in AI ethics in favor of making AI systems more transparent, and trying to combat the black-box approach of machine learning. Embedding ML into socio-technical governance - including in the principles of Algorithmic Impact Assessment (AIA) and Strategic Environmental Assessment (SEA) and through the presence of a minimum threshold for benefits to socially vulnerable groups (fairness constraint) (Baresi and Russo 2025). But parallel scholarship of green gentrification warns that if equity is not to be part of the logic of investment, then a vulnerable population can be marginalized from the way by the green investment (Anguelovski et al., 2022). Explicit negotiation between trade-offs of thermal comfort, ecological corridors, capital expenditure, spatial justice can be achieved with multi-objective reinforcement learning, as compared to single objective optimisation (e.g., maximise total green cover).

3. Materials and Methods

The proposed methodological pipeline comprises four core layers: Data Fusion, DRL Problem Formulation, Multi-Objective Reward Function Design, and Governance Interface Integration.

3.1 Data Fusion and 3D Urban Digital Twin Construction

The 3D urban canopy has been created by combining the surface models derived from UAV photo graphometry and the multi-spectral data of Sentinel-2, the cadastral data/building footprint data from Survey Office, Dang in Ghorahi Sub-Metropolitan City. The baseline vegetation typologies and buildings heights, wall surface areas and Sky View factors (SVF) are mapped by using Convolutional Neural Networks (CNN) and random forest classifiers. Spatial raster grid is generated by integrating socio-economic vulnerability indicator of NSO Nepal census 2021 with 20 m × 20 m grid resolution which is composed of multidimensional poverty indicator, age and pre-existing health conditions and thermal exposure indicators.

3.1.1 Case Study Area and Multi-Source Dataset Specification

This is a descriptive study that included women. In this study, an empirical study methodology was used with the women of Ghorahi Sub-Metropolitan city located in Dang district, Lumbini province (28° 03'N, 82° 49'E) of Nepal. It is situated in the valley of Inner Terai (Bhitri Madhesh) between Mahabharat Range in the north and Churia (Siwalik) Hill in the south and the corridor of Rapti River runs through it, making it one of the fastest growing cities of Nepal's mid-western region. The study area was chosen based on the need to have a continuous area of 5km2, which was located approximately around the CBD and the high density mixed residential areas of the city and the river corridor. The Zone has been having a high level of Intra-urban Heat Exposure for the past decade and ecological patchiness, due to high rate of land use changes, midrise construction and loss of green cover of Agro-ecological Areas and Riparian Areas. The average FAR of the core zone is 1.8 and the emerging clusters have more than 3.0. The fast-densifying valley is an in-between space between the thinly vegetated hill ecology and the density of the plain that is tested for optimization of 3D spatial green infrastructure, has a complex canopy geometry, a high rate of land-use changes and fragile microclimates.

The optimization model is developed by using multi source data fusion pipeline consisting of four basic spatial datasets which are helpful to the availability of these data at the local level in Ghorahi Sub-Metropolitan City:

Fine scale Digital Surface Models (DSM) and Digital Elevation Models (DEM) were generated as well as Canopy Height Models (CHM) for the post monsoon clear sky period of low altitude UAV flights over the 5 km2 urban testbeds.

Multi-seasonal 10 m multispectral image data from Sentinel-2 were used to create baseline Normalized Difference Vegetation Index (NDVI) and surface reflectance profiles for the Ghorahi study area in the summer months when thermal stress is at its peak.

All the cadastre layers, vectors building footprints were obtained from the Survey Office and the number of floors, ownership boundaries and public/private realm designation were gathered from the Dang and Ghorahi Sub-Metropolitan city (SMC) municipal office in the study area.

Social-economic indicators of social equity were developed by performing an aggregation of some of the indicators of the NSO Census 2021 on a raster-grid of 20m x 20m.

3.2 Deep Reinforcement Learning Formulation and Algorithmic Architecture

The discretized domain takes a space of 5 km2, at a spatial resolution of 20 m × 20 m, and has a number of decision cells N = 12,500. The spatial resolution of a decision cell is 20 m × 20 m and the number of cells is N = 12500. To make all these 12,500 grid cells computationally tractable, we use an approach of sharing the parameters across the network with agent-index embedding, which means that only one pair of actor-critic networks for each grid cell.

The network architecture is Multi-Agent Deep Deterministic Policy Gradient (MADDPG; Lowe et al., 2017) with the training being centralized and the execution decentralized. The actor network π_i has four fully connected layers (256, 256, 128, 64 units), with ReLU activation functions. The discrete action space is dealt with by introducing a

Gumbel-Softmax discretization layer with a parameter of Gumbel temperature that is linearly annealed from 2.0 to 0.1 during training for the policy, over the six categories of interventions. Instead of the typical Ornstein-Uhenbeck noise, the exploration policy is an annealed epsilon-greedy policy, which goes from 1.0 to 0.05. The centralized critic Q_i is a five-layer deep network having 512, 512, 256, 128 and 64 units respectively with the LeakyReLU activation functions (α=0.01).

State space (S): The grid cell has characteristics of Building density (3D), distribution of BGI (current), velocity of local air flow, solar exposure, heat vulnerability index and proximity to the ecological core.

Action Space (A): Discrete interventions in space per cell: (1) No Action, (2) Intensive Green Roof, (3) Extensive Green Roof, (4) Vertical Living Wall, (5) Street Tree Canopy Enhancement, (6) Permeable Pervious Bioswale.

A multi-objective reward function (R) is used to balance 3D microclimatic cooling (ΔT_UTCI), landscape ecological network connectivity (ΔC_eco) and algorithmic spatial equity (F_equity) with constraints on capital expenditure (C_inst), to obtain a composite reward scalar (R_total):

$$R_total = w_1 \cdot \Delta T_UTCI + w_2 \cdot \Delta C_eco + w_3 \cdot F_equity - w_4 \cdot C_inst$$

All four Pareto weighting coefficients w_1, w_2, w_3 and w_4 are numbers in the range [0, 1] with $w_1 + w_2 + w_3 + w_4 = 1$. The weights will be: w1 = 0.35, w2 = 0.25, w3 = 0.25, w4 = 0.15, for 3D thermal mitigation, landscape ecological connectivity, algorithmic spatial equity and installation and maintenance capital cost penalty. Just one sentence of explanation on how these weights have been obtained, such as expert elicitation or grid search.

The agent MADDPG can interact with several microclimate simulation environments (ENVI-met / InVEST integrations) at once and be able to find Pareto optimal spatial deployment configurations.

3.2.1 Hyperparameter Specifications and Convergence Criteria

For the MADDPG algorithm, the Adam optimizer hyper parameters are: actor learning rate = 1×10^{-4}, critic learning rate = 1×10^{-3}, discount factor γ = 0.95, target network soft update rate τ = 0.005, a replay buffer containing 1×10^{6} transitions, and a minibatch size of 128. As the actions are discrete, exploration is performed using Gumbel-Softmax policy with an annealed ε-greedy sampling policy that linearly decreases from 1.0 to 0.05, over 3,000 train steps. The episodes consist of 50 spatial placements each and training is 5,000 episodes. The 200-episode window convergence of the algorithm is defined when the average critic loss variance is lower than 0.001.The convergence of the algorithm is defined as 200 episodes window when the average critic loss variance is less than 0.001, and the Pareto front hypervolume metrics are stable. The algorithm's convergence is evaluated with a window size of 200 episodes and a relatively stable hypervolume measure of the Pareto front with a criterion of the average variance of the critic loss below 0.001. Add hardware information and overall wall clock time spent in training.

3.2.2 Microclimate and Eco-Hydrological Simulation Settings

A coupled 3D microclimate simulation environment (ENVI-met v5.1) coupled with a fast and physics-informed surrogate neural network trained on 1200 baseline ENVI-met simulations is used to calculate the microclimate reward feedback, ΔT_UTCI. The representative maximum atmospheric forcing conditions recorded by the Department of Hydrology and Meteorology (DHM) at Ghorahi, Dang station during peak summer conditions were used for setting up the model parameters of ENVI-met. Air temperature was 38.5 °C, relative humidity 35%, background wind speed at 10 m 2.1 m/s in SW direction (225°) which corresponds to the maximum conditions during the summer solstice. The soil heat flux and the thermal transmittance of the building walls (U-values) were assumed according to Nepal Building Code and municipal building code of Ghorahi Sub-Metropolitan City, $U_wall = 1.8\ W/m^2K$, $U_roof = 0.25\ W/m^2K$. The inVEST Urban Cooling module, the Morphological Spatial Pattern Analysis (MSPA; Soille & Vogt 2009) indicator and the Integral Index of Connectivity (IIC; Pascual-Hortal & Saura 2006) at a distance threshold of 500 m are used to measure ecological connectivity (ΔC_eco).

3.2.3 Microclimate Model Validation Protocols

The Ghorahi urban testbed was used to gather a set of empirical measurements of microclimate parameters at 18 fixed weather stations and pedestrian transects to evaluate the accuracy of microclimate surrogate model on the hottest days of thermal stress in summer. The validation results revealed acceptable agreement between the observed air temperature/UTCI and the simulated results (RMSE of 0.42 °C; MAPE of 1.18% and R^2 of 0.94) which are within acceptable range of model accuracy for microclimate modelling.

3.3 Socio-Technical Governance and Policy Alignment

The model for optimization explicitly supports fairness constraints to ensure principles of responsible data governance are followed, which are recognized as internationally agreed upon, and then adjusts them to the context of the Municipality stakeholders in Nepal without "automated spatial injustice." The optimized spatial scenarios are converted to policies score cards which are directly linked with the policies mentioned in municipal building bylaws, open space allocation standards, setback requirement and courtyard greening provision and compulsory building coverage ratio as mentioned in Local Government Operation Act (2017) in Ghorahi Sub-Metropolitan City. It's a local level regulatory convergence, a working mechanism for local municipalities to set up a green factor scoring and enforcement mechanism in Nepal.

4. Results

The three different allocation strategies are tested on 30 independent trainings with 3 random seeds initializations of the 5 km^2 urban testbed and 3 runs per allocation strategy for the empirical benchmark evaluation. These values are means over 5000 episodes and uncertainty bounds (Appendix, ± 1 standard deviation over 30 runs) are presented here. To ensure that the performance improvements that were observed from Strategy Cover Strategy A and B was statistically significant for all metrics ($p < 0.001$), a Wilcoxon signed rank test was also performed.

Table 1. Performance comparison of GI spatial optimization strategies.

Performance Indicator	Baseline (A)	2D Heuristic (B)	Proposed 3D DRL (C)
Mean pedestrian UTCI reduction (°C)	0.0 ± 0.0	−1.2 ± 0.18	−2.8 ± 0.23*
3D shading and canopy efficiency (%)	18.4 ± 1.1	32.1 ± 2.4	58.6 ± 3.1*
Ecological Connectivity Index (IIC)	0.14 ± 0.01	0.28 ± 0.03	0.45 ± 0.04*
High-vulnerability group coverage (%)	21.0 ± 1.5	38.5 ± 3.2	81.2 ± 4.6*

These results were found to be statistically significant with Strategies A and B (Wilcoxon signed rank test, $p < 0.001$, 30 independent runs).

As seen in Table 1, the performance of the proposed strategy (C) is much superior with respect to the traditional strategies. The factor of building height variation, solar exposure angles and wind corridors, and optimized for vertical greening, were included in the 3D DRL agent, which was able to reduce heat stress of the street-tree canopies in narrow urban street canyons by 2.8 ± 0.23 °C on average (133% reduction), compared to the 2D heuristic. The coverage of BGI interventions in high vulnerability neighborhoods was more than doubled with the embedding of the socio-economic fairness reward (81.2 ± 4.6%) compared to 2D (38.5 ± 3.2%). The models were found to be highly predictive about the measurements made by the field weather stations – with an RMSE of 0.42 °C, a MAPE of 1.18% and an R^2 of 0.94.

Sensitivity analysis of Pareto weights (w_1 – w_4) was carried out for various policy priorities. For the same policy scenario, an optimal BGI coverage increases from 52.4% to 89.1% obtained by an increase in equity weight (w_3) from 0.10 to 0.45 with an only slight decrease of 0.3 °C in the overall thermal reduction (UTCI reduction).

5. Discussion

The experimental results confirm the hypothesis that 3D morphological parameters are needed when planning a BGI in a complex urban environment with a vertical dimension. The 2D models created, using the conventional approach, could not effectively distribute the resources, since they only took into account the available space and the mean surface temperature, and did not consider the microclimate of the pedestrian level and the solar radiation received by the wall.

We concluded that the socio-technical approach to governance of Baresi and Russo 2025 still holds merit; that models of AI can be used as tools of decision support in socio-technical governance; but that models of AI cannot be used as a decision maker. The algorithmic optimization maps can be immediately mapped into enforceable building bylaws and developers' incentives, public investments, with the combination of procedures for Strategic Environmental Assessment (SEA) and municipal green-coverage development controls such as Melbourne Green Factor Tool (Bush et al., 2021).

5.1 Limitations and Future Research

The proposed framework is a first step towards optimization of 3D BGI, but there are some limitations. On larger urban scales the use of RL training is challenging since the coupled 3D microclimate simulations are very costly to compute. The trained policy has not yet been compared to the validity to other cities that have different topographical and climatic conditions with the Ghorahi testbed in Inner Terai valley. Evaluation in the field of the models is also necessary to see if the desired microclimate and connectivity benefits are achieved in the longer-term (multi-year) with tangible, public health and ecological outcomes. In future, use of a physics informed surrogate model will help speed up the training process, extension of the proposed architecture to include participatory sensing data collected longitudinally will be extended to other municipalities undergoing rapid densification in Nepal as per the Local Government Operation Act 2017 will be encouraged.

6. Conclusion

A multi-objective Deep Reinforcement Learning (DRL) framework is proposed and exemplifies it with a Sub-Metropolitan City Ghorahi (SMC) of Dang, Nepal Himalayas with a high rate of urbanization. The framework seeks to address the issue of spatial justice and urban AI modelling that has been missing in the past, by bringing together the 3D bioclimatic metrics, landscape ecological connectivity, algorithmic fairness and development control governance mechanisms. The results indicate that the 3D urban morphology modelling and the socio-economic vulnerability of people results in a much higher thermal comfort efficiency and a more balanced resources distribution. The methodology offers a transferable process for shifting towards a more resilient, sustainable and equitable urbanism, from one of rapid urban densification, that is vulnerable to climate change.

Declarations

Funding: This research received no external funding.
Conflicts of interest: The authors declare no conflict of interest.
Acknowledgments: Thank Ghorahi Sub-Metropolitan City, the Survey Office Dang, NSO, DHM, and field assistants as applicable.
ORCID iD: **Er. Bikram Babu KC Orcid:** 0009-0008-2060-1395
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers
Artificial Intelligence (AI) Use Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

References

[1] Ahern, J. (2007). Green infrastructure for cities: The spatial dimension. Cities and the Environment, 1(1), 1–14.
[2] Anguelovski, I., Connolly, J. J. T., Cole, H., Garcia-Lamarca, M., Triguero-Mas, M., Baró, F., Martin, N., Conesa, D., Shokry, G., del Pulgar, C. P., Argüelles, L., & Pérez Del Pulgar, C. (2022). Green gentrification in European and North American cities. Nature Communications, 13, 3816.
[3] Baresi, U., & Russo, A. (2025). Can machine learning support planning for equitable green infrastructure? A perspective on opportunities, risks, and ethical pathways. Urban Science, [volume, issue, pages verify].
[4] Bush, J., Ashley, G., Foster, B., & Hall, G. (2021). Integrating green infrastructure into urban planning: Developing Melbourne's Green Factor Tool. Urban Planning, 6(1), 28–40.
[5] Jendritzky, G., de Dear, R., & Havenith, G. (2012). UTCI Referencing the human thermal environment for urban climate applications. International Journal of Biometeorology, 56(3), 481–494.
[6] Lafortezza, R., Davies, C., Sanesi, G., & Konijnendijk, C. C. (2013). Green Infrastructure as a tool to support spatial planning in European urban regions. iForest – Biogeosciences and Forestry, 6, 102–108.
[7] Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. Advances in Neural Information Processing Systems, 30, 6379–6390.
[8] Ma, X., Ye, J., Yang, F.-S., Tang, S., & Jiang, Z. (2025). Machine learning application in investigating cooling effect of urban blue–green infrastructure: A systematic review. Technologies, [volume, issue, pages verify].
[9] Milojevic-Dupont, N., & Creutzig, F. (2019). Low-carbon urban planning with machine learning. Proceedings of the 36th International Conference on Machine Learning (ICML), PMLR 97, [pages].
[10] Pascual-Hortal, L., & Saura, S. (2006). Comparison and development of new graph-based landscape connectivity indices: Towards the priorization of habitat patches and corridors for conservation. Landscape Ecology, 21(7), 959–967.
[11] Soille, P., & Vogt, P. (2009). Morphological spatial pattern analysis. International Journal of Remote Sensing, 30(6), 1667–1683.
[12] Department of Hydrology and Meteorology (DHM), Government of Nepal. Climatological records for Dang region.
[13] National Statistics Office (NSO), Government of Nepal. (2021). National Population and Housing Census 2021. Kathmandu: NSO.
[14] Yeo, O. T. S. (2024). Frameworks of ecological connectivity: Systematic review on green infrastructure's case studies. GeoJournal, 89(4).