Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multi objective Optimization for Sustainable Urban Land-Use Planning Under Ecological Constraints

Cui Jing
Hebei University of Engineering Science
Yiyun Wang
Hebei University of Engineering Science
Lina Zhao
Linazhao178@gmail.com

Hebei University of Engineering Science

Research Article

Keywords: Urban land-use planning, multi-objective optimization, Urban sustainability, Ecological constraints, Feature sensitivity analysis, Evolutionary optimization

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11505560/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Multi objective Optimization for Sustainable Urban Land-Use Planning Under Ecological Constraints

Cui Jing[1], Yiyun Wang[1], Lina Zhao [1*]

1. College of Civil Engineering and Architecture, Hebei University of Engineering Science, Shijiazhuang 050091, Hebei, China

*Linazhao178@gmail.com

Abstract

Achieving sustainability and environmental quality, public safety and economic development are complex trade-offs in urban land-use planning. These goals are typically planned separately, which reduces their ability to put together solutions that are both balanced and sustainable. This study suggests a framework for multi-objective optimisation, which combines data-driven predictive modelling, feature sensitivity analysis and evolutionary optimization, for sustainable urban land use. A surrogate model called A Histogram Based Gradient Boosting (HistGB) is created to explore the nonlinear relationships between 15 sustainability indicators, and the hyper parameters of HistGB are optimized by Novel Hybrid Kepler Optimization Algorithm (NOHKOA). The Dynamic Trade-off Coefficient Selection Multi-Objective (DTCSMO) algorithm is then used to optimize conflicting sustainability, environmental burden and urban risk goals while integrating continuous and discrete planning variables. The coefficient of determination for the test data reaches 0.9877 and the mean absolute error (MAE) is 0.0152 and the root mean squared error (RMSE) is 0.0193 for the HistGB model. For feature sensitivity analysis, the most sensitive features are green cover (mean change in coefficient of determination: 0.8196) and renewable energy use (mean change in coefficient of determination: 0.3584). The results of the ablation experiment also show the significance of the environmental indicator, and the test coefficient of determination drops to 0.1542 when the environmental indicator is removed. The DTCSMO algorithm can produce 120 Pareto-optimal solutions, allowing decision makers to consider different trade-offs between the planning goals. For the evaluation of the characteristics of the solutions a comparison with the multi-objective particle swarm optimization algorithm is additionally provided. In summary, the proposed framework provides a data-driven method for

urban land-use planning that is accurate, interprets the data through sensitivity analysis and incorporates multiple objectives in decision making.

Keywords: Urban land-use planning; multi-objective optimization; Urban sustainability; Ecological constraints; Feature sensitivity analysis; Evolutionary optimization.

1. Introduction

1.1.Background

The land resources are important in the social and economic development. Sustainability and resilience have been found to be major subjects of the rational utilization of land resources. The land use allocation is becoming very significant accompanying the urbanization process and the economic growth, and has gained significant interest. The land use allocation problem in context of sustainable development is a problem that allocates various competitive land uses to spatial units with the aim of achieving the objectives of land managers that are desired [1]. In the context of the dual stressors of increased human activity and global warming, land use spatial optimization has been in the limelight of geographical studies. Irrational land use patterns may cause serious environmental issues and problems with resources and energy use, as well as an imposing impact on the sustainable development of cities [2]. As human activities grow in intensity and the earth undergoes an irreversible change with a rapid socio-economic development, the ecosystem services (ES) are gradually becoming degraded and ecological risk (ER) elevates considerably. The change in land use is one of the key drivers that affect the evolution of the regional ecosystem due to its ability to change physical, chemical and biological properties of the land surface, and also changes in hydrological factors like precipitation, evaporation, and surface runoff [3]. The changes do not only alter the ecosystem structure and functioning, but they also influence the carbon and nitrogen cycles among other material and energy interactions [4]. ES within ecologically delicate regions is generally vulnerable to climatic changes and human disruptions. Thus, it is of high practical importance to minimize ER in the aforementioned regions by using optimized land use patterns [5]. Land use changes are also central to the carbon capture and emissions. Optimized land use configurations as noted by the Intergovernmental Panel on Climate Change (IPCC) has tremendous potential of reduction in carbon emissions. It has been proved that by modifying the land use patterns, it is possible to decrease transportation demand, energy consumption as well as increase the land use efficiency which is fundamental in terms of meeting

energy conservation and emission reduction targets [6]. Urban agglomerations have become new geographical units that contribute to the cross-administrative cooperation and to the increase of national competitiveness in the division of labor in the world. Nevertheless, most of the urban agglomerations that were formed at an early age, as well as the most mature ones, have excessively focused on the economic progression, thus leading to the uncontrolled growth of construction lands and uneven distributions of land use patterns without attention to the ecological protection. This has contributed towards severe urban diseases, such as the lack of resources, environmental pollution, and ecological degradation [7]. Previous literature has demonstrated that land use optimization can be successful to improve spatial allocation to high-quality development through providing better patch contiguity, economic performance, and ecological utility within the current planning regulations and policy frameworks [8].

1.2.Literature review

Recent researches show that the concept of multi-objective optimization of urban land use with ecological constraints has become one of the focal research topics in sustainable spatial planning. The use of a multi-objective problem can serve as an example since Tingting Pan (2023) used a genetic algorithm that is based on vectors in order to tackle the multi-functional land-use zoning. The optimization framework with simultaneous consideration of the ES value, economic efficiency, and spatial compactness is selected in the study, a strategy that proved that evolutionary methods can play a significant role in enhancing coordinated spatial allocation [8]. Yuxiang Dong et al. (2026) suggested a multi-objective optimization model, in which neural networks-based surrogate models U-shaped Network (UNet) and Attention UNet are utilized to estimate ES in urban green infrastructure design. This framework was at that point incorporated into a multi-objective evolutionary optimization procedure in order to decide on the optimum allocation of lands to achieve maximum ES as well as minimizing the conversion of land which highly reduced the computational time [9]. Based on the socio-ecological systems approach, Rodriguez-Gonzalez (2025) formulated the multi-objective optimization model that will help to reinforce the resilience of the region and reduce the risk of climate change. The results highlighted that land-use structure plays a very essential role in mitigating climate-related hazards [10]. Cheng Qianwen (2026) integrated NSGA-II and spatial simulation model to optimize the ecological security and land-use suitability in the regional scale. Predictive spatial modeling in conjunction with the evolutionary optimization facilitated the accuracy of decision-making and the ability to analyze scenarios [11].

Also, Sabab Aosaf (2025) suggested highly developed computational intelligence strategies in land allocation of multiple objectives in complex urban areas, pointing to the faster convergence speed, increased search efficiency, and more effective resolution of conflicting goals [12]. In a study by Yian Wang, Xueming Zhou and Kai Cao (2026), they have designed a hybrid multi-objective optimization framework where they used NSGA-III and Differential Evolution in improving the livability of urban areas. Their study revealed that spatial optimization has the ability to optimize economic development, environmental sustainability and social equity which presents a better-rounded approach to urban land-use planning considering conflicting goals [13]. Rui Shu (2024) proposed a multi-scenario land-use optimization model of the fine scale land survey data, combining the spatial simulation and ES value estimation. The paper has compared the trade-offs in terms of development and ecological protection, and the results indicated that the optimal development of land-use patterns can improve ES and sustainability in urban development [14]. In the article by Cruz (2024), surrogate-based multi-objective optimization of sustainable building design was reviewed, and it is noted that the combination of machine learning models and evolutionary algorithms can help to fasten evaluation and assist in trade-off analysis [15]. The article by Shan (2025) has examined AI-based multi-objective optimization in the urban energy retrofit context, focusing on application as machine learning surrogate models to lower the computational cost and enhance decision-making under competing sustainability goals [16].

1.3.Novelty

The main novelty of this study is the simultaneous use of surrogate-based predictive modeling, feature sensitivity analysis, and multi-objective evolutionary optimization, and combining them in a single framework for sustainable urban land-use planning. The key features are:

1. Integrated data-driven planning framework: Using a novel hybrid Kepler optimization algorithm for hyperparameter optimization, a surrogate model, namely HistGB, is combined to model the 15 urban sustainability indicators while considering their nonlinear relationships.
2. Adaptive multi-objective optimization: The Dynamic Trade-off Coefficient Selection Multi-Objective algorithm is used to optimize urban sustainability, environmental burden and urban risk under the consideration of continuous and discrete planning variables.
3. Quantitative identification of critical sustainability indicators: Feature sensitivity and ablation analysis reveal that green cover and renewable energy consumption are important features that

affect prediction performance and show that the environment has a significant impact on the predictive results.

4. Pareto-based decision support: The proposed framework returns 120 Pareto-optimal solutions for planners to consider, instead of one optimized land use set of configurations.
5. Comparative evaluation: The optimization results are compared with those obtained by multi-objective particle swarm optimization to see the differences in the objective values and the resulting land-use configurations.
6. All of these contributions combine to give an intelligible and adaptable decision-support system to assess competing sustainability goals within urban land-use planning processes.

2. Methodology

Fig.1 shows that the multi-objective optimization of urban land-use planning, considering the environmental constraints, actually occurs in the following steps: In Step 1- Data Collection, data is gathered from sustainable urban planning data sets, and all data is normalized without any loss of data. Step 2 - Surrogate Model Development, surrogate model is developed using HistGB algorithm, which reflects hidden properties of the data. Step 3 Hyperparameter tuning entails optimum hyperparameters on the model with the help of the NOHKOA. The use of the Permutation Importance method to understand feature importance of the features potentially important to the end product (e.g. green coverage and renewable energy) will be utilized in Step 4 – Feature Sensitivity Analysis. Ablation Study (Step 5) is a study that performs to investigate the feature groups that have a critical impact and that have a low influence on land-use outcomes. Then in Step 6 of Multi-Objective Optimization the DTCSMO algorithm is implemented to optimize the objectives of maximizing sustainability, minimizing load and risk, and the Pareto-optimal sets of the best possible combinations are obtained.

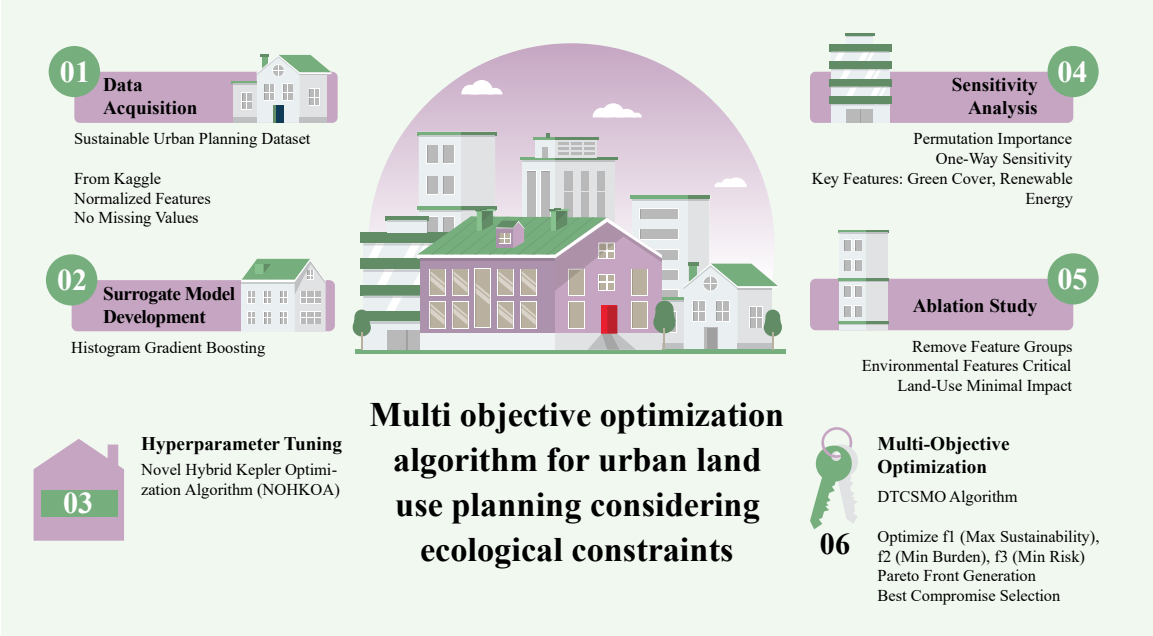


Fig.1. Sequential procedure of the multi-objective optimization model of the urban land-use planning under environmental restrictions.

2.1.Data Collection

The dataset applied in the given study is the Sustainable Urban Planning and Landscape Dataset [17]. The characteristics are environmental indicators, social indicators, infrastructural indicators, a risk index, and a land-use category coded by one-hot representation. All values are brought in [0,1] range. The dataset does not have any missing data or one-hot coding errors.

2.2.Surrogate Model Development

HistGB is an ensemble learning algorithm in the form of decision trees, which enhances the speed and computation efficiency through higher-order continuous features by quantifying them into histogram bins. The model is built in a step-by-step fashion such that each step a new decision tree is added to learn the negative gradient of the loss function with respect to the existing model prediction as in Eq. (1):

$$F_t(x) = F_{t-1}(x) + \eta h_t(x) \tag{1}$$

where $F_m(x)$denotes the model output at iteration m, $h_m(x)$represents the decision tree trained on the negative gradient of the loss function, and $\eta$is the learning rate controlling the contribution of each tree. The computation of the splits is done in terms of discretized feature bins in the histogram-based form, a very important factor that greatly decreases the computational cost but does not compromise predictive accuracy [18].

HistGB model is created in order to forecast the score of urban sustainability. Continuous values are quantized into histogram bins and are learnt with decision-tree structures. The model serves as a proxy that estimates computationally intensive estimates of objective functions. The data is divided into training and testing sets and the cross-validation (CV) is used to ascertain the model stability.

2.3.Hyperparameter Tuning by the Novel Hybrid Kepler Optimization Algorithm (NOHKOA)

NOHKOA is an original hybrid optimization algorithm that is based on Kepler laws, and is created to be able to search hyperparameter space of machine learning models efficiently. It is a hybrid of international exploration and national exploitation to find the best hyperparameter estimates that can optimize the model accuracy and reduction of prediction error [19].

Fig.2 gives the detailed description of an integrated framework that is composed of HistGB surrogate model and two evolutionary optimization algorithms, DTCSMO and NOHKOA. The process of the HistGB model is depicted in the upper-left part. The continuous features are discretized and sorted into histogram bins first followed by the learning process which is done in a decision-tree structure. This component is a surrogate model, which has a high predictive accuracy of the objective function value (e.g., urban sustainability index) and is used in place of computationally expensive evaluations. The flowchart of the DTCSMO algorithm is provided in the bottom-left end. The algorithm starts with the population initiation, fitness of individuals (starlings) is calculated, and the algorithm uses vertical crossover operator and horizontal crossover operator, population division, separation strategy, multi-flock formation and Dynamic Opposition to improve variety and prevent local optima. The quality of flocks is evaluated and the global optimum updated in every iteration until the stopping criterion is met. The flowchart of the NOHKOA algorithm which is inspired by gravitational interactions is presented on the right side. The algorithm begins by initializing the population and determining the global best solution followed by the calculation of Euclidean distance, gravitation force and particle velocity. Candidate solutions are updated using random numbers and a conditional update mechanism and the best-so-far solution is kept because, as soon as the maximum number of iterations is met, candidate solutions are not updated. In general, the figure indicates that the HistGB model is an effective predictor in addition to the evolutionary algorithms DTCSMO and NOHKOA that allow

exploring the solution space faster, more stable, and efficient. This hybrid is a machine learning based optimization model of solving multi-objective problems.

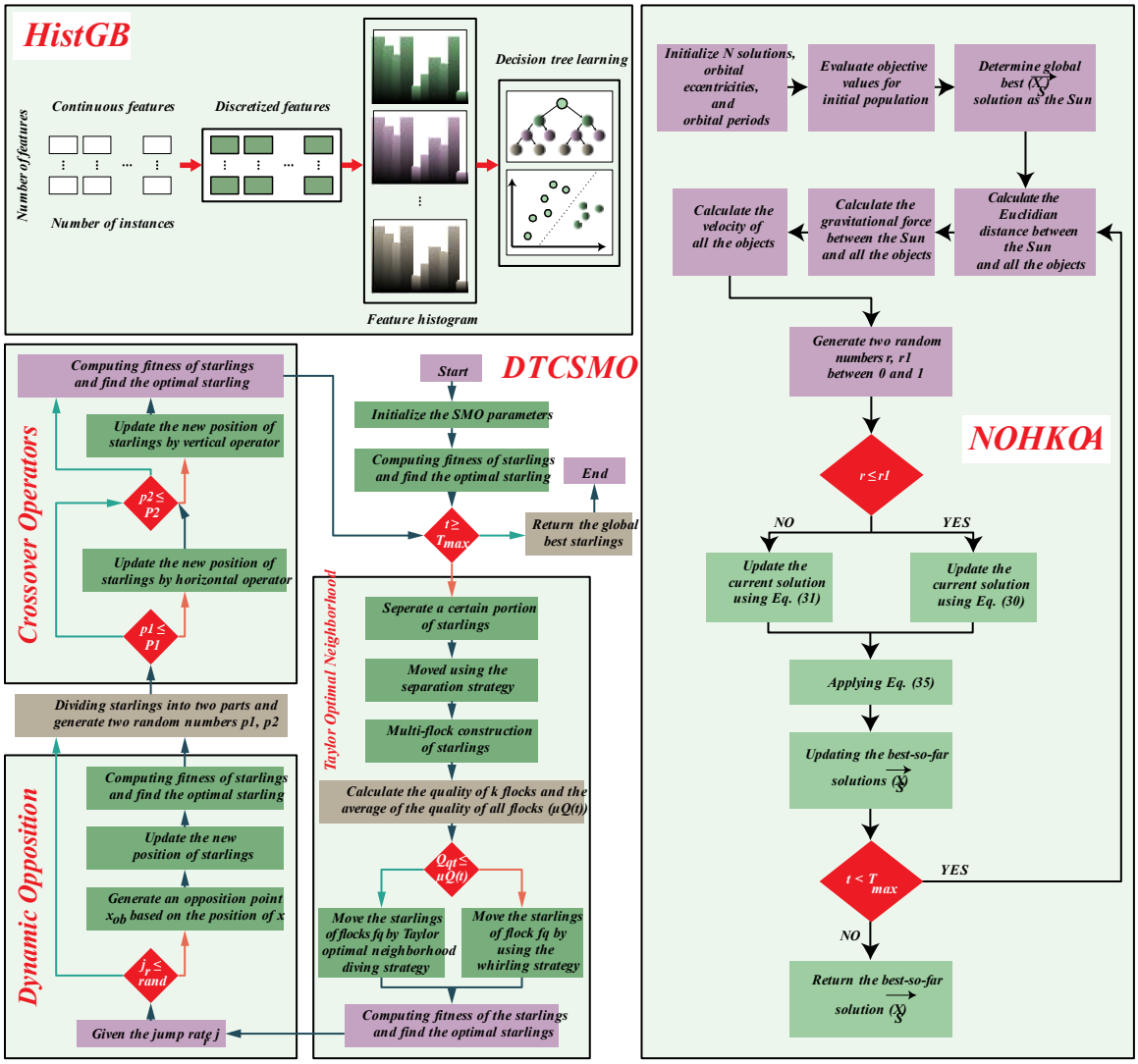


Fig.2. Hybrid surrogate-assisted multi-objective optimization model that incorporates HistGB, DTCSMO, and NOHKOA.

2.4.Sensitivity Analysis

Sensitivity analysis is used to determine the effect of the input indicators using:

- **Permutation Importance:** It involves permutation of every feature separately and the resultant variation in the model predictions is measured to determine influential features.
- **One-Way Sensitivity Analysis:** This method assumes the systematic variation of all variables within their range with the other variables held constant to be able to see the impact of a particular variable on the predicted outcomes.

2.5.Ablation Study

This model is retrained several times through systematic elimination of one group of features at a time. This method assesses the role played by each of the categories such as environmental, socioeconomic, risk and land-use indicators on model behavior.

2.6.Multi-Objective Optimization by the Dynamic Trade-off Coefficient Selection Multi-Objective (DTCSMO)

DTCSMO is a multi-objective optimization approach which does not utilize fixed weights but dynamically changes trade-off coefficients. This will enable the algorithm to strike a compromise of conflicting goals in the search process [20]. An optimization problem with three objectives was developed to achieve a balance of sustainability performance, environmental impact, and risk pressure on the urban areas. According to Table 1, the first objective (f1) is one that maximizes sustainability as it reduces the negative forecasted score. The second goal (f2) ensures the reduction of environmental burden through the minimization of carbon emission and air pollution. The third goal (f3) reduces risk pressure through lowering the risk of disaster, crime, and density and enhancing renewable energy and green cover.

Table 1. Developed multi-objective functions of sustainability maximization.

Label	Objective	Formula	Description
f1	Max Sustainability	− predicted_ score	Maximize sustainability
f2	Min Environmental Burden	carbon + air	Reduce pollution
f3	Min Risk-Pressure	risk + crime + density − (renewable + green)	Improve safety and resilience

3. Results

3.1.Predictive Model: Histogram Gradient Boosting

HistGB model was also trained to predict the urban_ sustainability_ score. Table 2 presents performance results that demonstrate that there is stable and precise performance in cross-validation and the test set. The low standard deviation in CV will also mean that the model is stable and the fact that the difference between CV and test results is minimal will be used to show that there is indeed no overfitting.

Table 2. Predictive performance of the HistGB surrogate model during CV and test evaluation.

Metric	CV Mean ± Std	Test Set	Interpretation
R^2	0.9847 ± 0.0006	0.9877	Model accuracy
MAE	0.0165 ± 0.0002	0.0152	Mean Absolute Error
RMSE	0.0208 ± 0.0004	0.0193	Root Mean Squared Error

Table 3 summarizes the model hyper parameters. To control the complexity and overfitting, the model was trained with the moderate minimum tree depth and low learning rate. The train / test split, CV folds, iteration and random seed are also reported in the table.

Table 3. Hyperparameter set of the HistGB surrogate model.

Parameter	Value	Notes
Algorithm	HistGB	sklearn
Max depth	6	Complexity control
Learning rate	0.05	Gradual learning
Max iterations	500	Number of trees
CV folds	5	Cross-validation
Train/Test	80/20	Data split
Random seed	42	Reproducibility

3.2.Sensitivity Analysis

Each feature was randomly permuted and the corresponding decrease in $R^2$was measured. As shown in Table 4, removing green_ cover_ percentage nearly eliminates the model's predictive power ($\Delta R^2 = 0.8196$), indicating its dominant role. renewable_ energy_ usage ranks second in importance. disaster_ risk_ index and carbon_ footprint show moderately high influence, while the remaining indicators have limited impact. In this stage, each variable was independently varied from 0 to 1 while keeping other variables constant.

Table 4. Permutation feature importance results with effect on R^2.

Feature	Mean ΔR^2	Std	Influence
green_ cover_ percentage	0.8196	±0.0379	Very High
renewable_ energy_ usage	0.3584	±0.0101	High
disaster_ risk_ index	0.2039	±0.0085	Moderately High
carbon_ footprint	0.1997	±0.0090	Moderately High
public_ transport_ access	0.0812	±0.0031	Moderate
crime_ rate	0.0775	±0.0028	Moderate
Other indicators	≈0	—	Negligible

The Permutation Feature Importance outcome of top 15 features in terms of 30 repetitions is given in Fig.3. The heights of the pyramids show the average change in model performance (likely ΔR^2) following the permutation of that feature, and the error bars shown in red are the standard deviation. Green_ cover is the most critical among the features, it leads to the greatest drop in performance and it has a dominating presence in the model structure. The second contributor to predictive power is renewable energy that is followed by the renewable_ energy. Following these two indicators, some features have a moderate implication with significantly lesser effects. When the ranking is beyond the sixth rank, importance of features becomes drastically low and tends to reach values close to zero meaning that it does not contribute much to the performance of models.

The error bars are generally small, indicating that the estimates of importance are generally constant and unstable. All in all, the distribution indicates a concentrated form whereby few key indicators result in most of the predictive ability of the model.

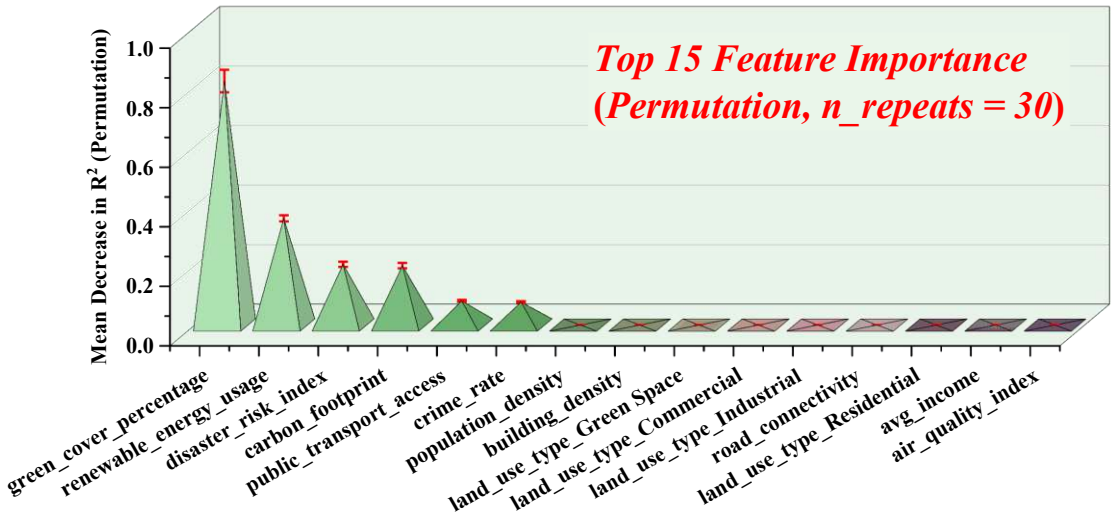


Fig.3. Importance of the best 15 indicators based on permutation feature importance.

According to Table 5, green_ cover produces the largest variation in the sustainability score (Δ = 0.370). renewable_ energy follows with Δ = 0.244. The remaining variables show smaller effects.

Table 5. One-way sensitivity analysis with changes in sustainability score.

Feature	Sweep	Min	Max	Δ
green_ cover	0→1	0.293	0.663	0.370
renewable_ energy	0→1	0.386	0.630	0.244
disaster_ risk	0→1	0.386	0.561	0.175
carbon	0→1	0.410	0.553	0.143
Others	—	—	—	Very Low

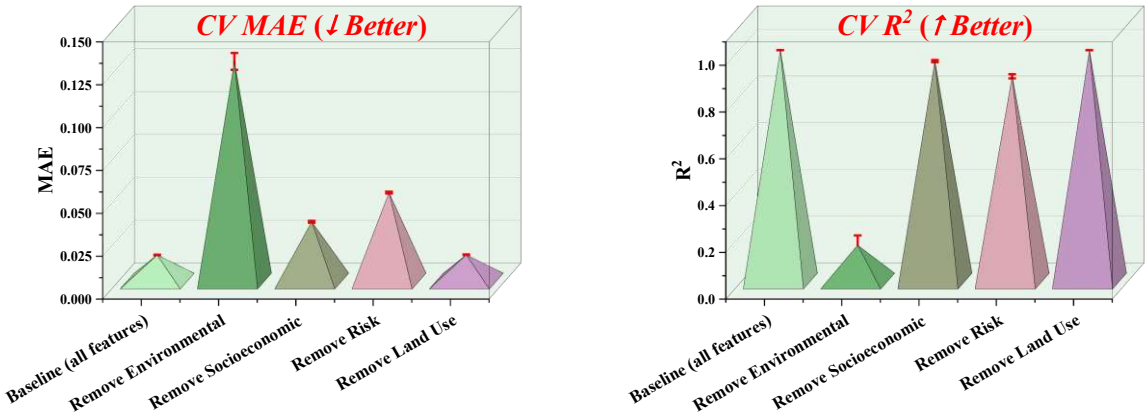
3.3.Feature Group Ablation Study

The model was retrained to evaluate the amount of contribution of each feature group by sequentially eliminating a category. By taking out the environmental indicators as Table 6 reveals, there is a drastic decrease in the performance (Test R^2 decreases from 0.9877 to 0.1542), which validates the fact that they are primarily involved in the model. The elimination of socioeconomic indicators as well as risk indicators result in moderate accuracy drops. On the other hand, elimination of land-use indicators has practically no impact on performance, signifying an insignificant contribution.

Table 6. Performance changes with a change in the feature group removal are present in the Ablation study results.

Configuration	#Features	CV R²	Test R²	CV MAE	Test RMSE
Baseline	15	0.9847	0.9877	0.0165	0.0193
Remove Environmental	11	0.0605	0.1542	0.1328	0.1601
Remove Socioeconomic	13	0.9323	0.9407	0.0361	0.0424
Remove Risk	14	0.8608	0.8759	0.0522	0.0613
Remove Land-Use	11	0.9848	0.9876	0.0164	0.0194

Three plots in Fig.4 depict the model performance in three metrics in case of CV. The Baseline and Remove Land-Use models perform better than the other ones since in the CV MAE (↓ Better) plot they have the lowest errors. On the contrary, the lowest performance and high error rate are shown by the Remove Environmental model. The Remove Social economic and the Remove Risk models lie between the extreme ends. The Baseline and Remove Land-Use models in the CV R^2 (↑ Better) plot also depict the greatest explained variance. They are next Remove Socioeconomic, then Remove Risk, and lastly, Remove Environmental has the lowest R^2. The CV RMSE (↓ Better) plot is no different: Baseline and Remove Land-Use are the best to use, Remove Socioeconomic and Remove Risk are mediocre, and the worst is Remove Environmental. In general, the Baseline and Remove Land-Use models are the most successful in their performance, and the removal of environmental indicators results in the most significant decrease in performance.



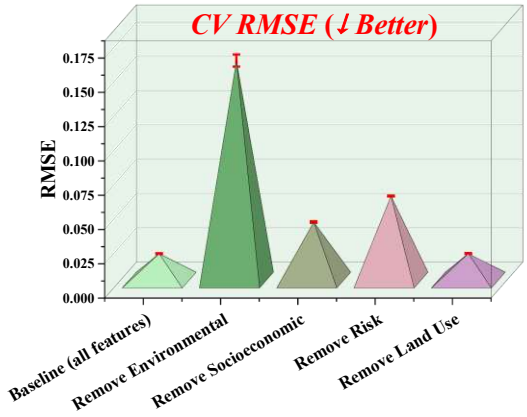


Fig.4. The plots of the three metrics are used to compare the performance of the model: CV MAE, CV R2 and CV RMSE.

3.4.Multi-Objective Optimization with DTCSMO

The optimization was carried out with the help of DTCSMO with the following parameters set as in Table 7. The algorithm was set to use 120 and 200 generations with the population size of 120. The balance between exploration and exploitation was achieved using a crossover rate of 0.90 and mutation index (η) of 20. The probability of mutation in land-use was 0.15. The total evaluations amounted to approximately 48,000 and the average time spent was 367 seconds.

Table 7. Parameters of DTCSMO algorithm applied in optimization.

Parameter	Value
Population	120
Generations	200
Crossover	0.90
Mutation η	20
Land-use mutation	0.15
Evaluations	~48,000
Runtime	~367 sec

The DTCSMO algorithm was able to produce 120 Pareto-optimal solutions which are various trade-off configurations among the three objectives. Table 8 indicates that the best-compromise solution was chosen using objective normalization. DTCSMO scored a little higher (1.0192) in Sustainability Score than MOPSO (1.0185). As MOPSO achieved slightly lower f2 and f3, DTCSMO achieved a more objective trade off. Both algorithms were drawn to almost maximum green cover and renewable energy (≈100%) in the best setup. Nonetheless, DTCSMO preserved small and yet attainable amounts of carbon footprint, crime rate, and population density, and MOPSO went to the extreme of zero in a variety of indicators. These findings imply that DTCSMO

provides a more balanced and practically viable compromise solution and at the same time the sustainability performance is very high.

Table 8. Best-compromise solution that was obtained by DTCSMO through a normalized objective selection.

Attribute	DTCSMO	MOPSO
Sustainability Score	1.0192	1.0185
f2	0.0242	0.0000
f3	−1.9803	−2.0000
Land-Use	Commercial	Residential
Green Cover	~100%	100%
Renewable Energy	100%	100%
Carbon Footprint	0.002	0.000
Crime Rate	0.003	0.000
Population Density	0.017	0.000
Disaster Risk	~0	0.000

Fig.5 shows the change of eight important indicators with the optimization or solution-ranking process. The patterns are the directional adjustment of variables in the multi-objective framework. The trends of the green_ cover_ percentage and renewable_ energy_ usage indicators depict an upward trend, which implies a gradual increase in the level of green infrastructure and integration of renewable energy. Likewise, access to public _transport is on an increasing trend as is the case with the sustainability-related attributes. Conversely, the trends of disaster_ risk_ index, carbon_ footprint, and crime_ rate are decreasing as it represents a decrease in both the burden of the environment and the risk of social aspects in all the optimized solutions. The avg_ income and population_ density indicators are relatively constant with a few changes meaning that they play a less sensitive or moderate role as opposed to the environmental and risk-related factors. In general, this number indicates that there is coordinated change in this sense that positive indicators are enhanced and negative ones reduced in the multi-objective optimization space.

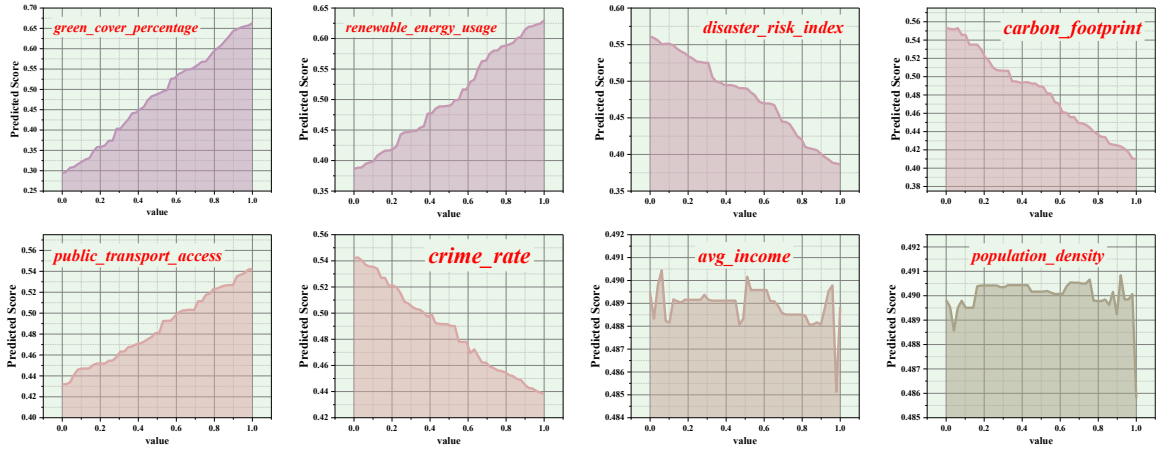


Fig.5. Trends of key sustainability indicators across optimized solutions.

Fig.6 shows a three-dimensional Pareto front which was obtained as a result of the multi-objective optimization process. The points indicate every possible trade-off between the three objective functions and the spatial distribution of points shows the nature of conflicts between them. The concentration of solutions in particular areas implies that the combinations of objectives that are more balanced are more achievable. The solution that is selected by the TOPSIS method is represented by the red-marked point that has coordinates (−1.008, 0.001, −1.999). This point is on the Pareto front boundary, and it is a compromise solution. The location of it indicates a moderate distance to the ideal and anti-ideal solutions in the objective space. In general, this figure shows the variety of Pareto-optimal solutions and the choice of a representative trade-off configuration in the framework of three objectives.

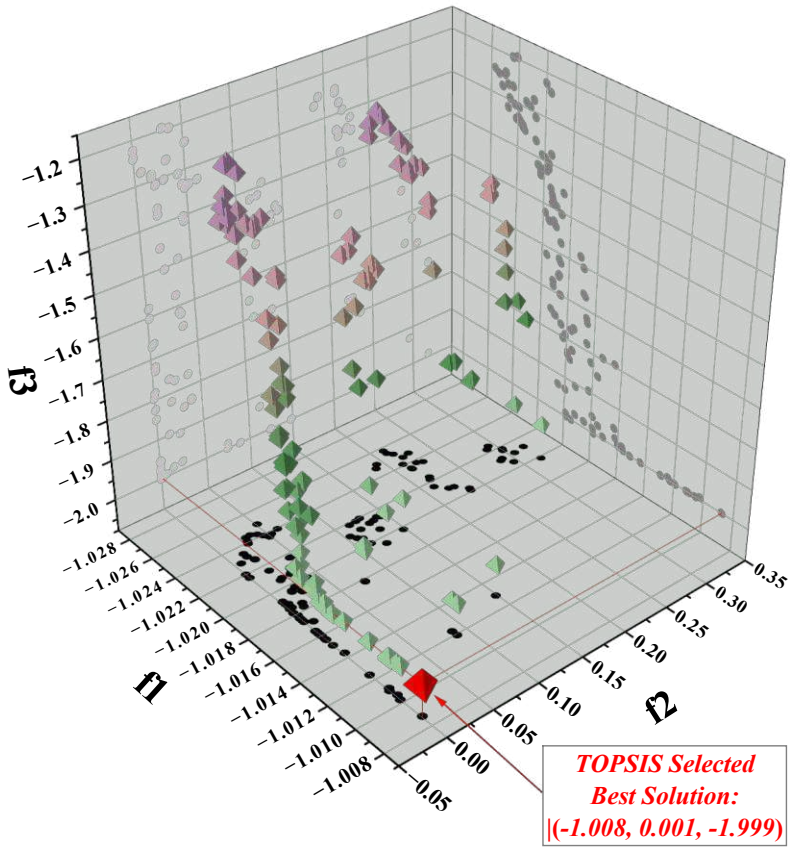


Fig.6. Three-dimensional pareto front having TOPSIS-screened compromise solution.

3.5.Graphical Analysis and Test Set Performance

The performance of the model on the test data is shown in Fig.7. The Test R^2 of 0.988 in the scatter plot presents very high predictive accuracy. The data points are tightly grouped along the 45-degree

reference line indicating that there exist high levels of agreement between the actual and predicted values with a low level of dispersion and no perceived systematic bias. The error distribution plot had all the residuals to be distributed normally and centered around 0, and the mean error is very near to zero. This implies that there is a randomness in the errors and that the model is statistically stable. On the whole, the findings support the predictive ability, adequate generalization capabilities, and confident model behavior.

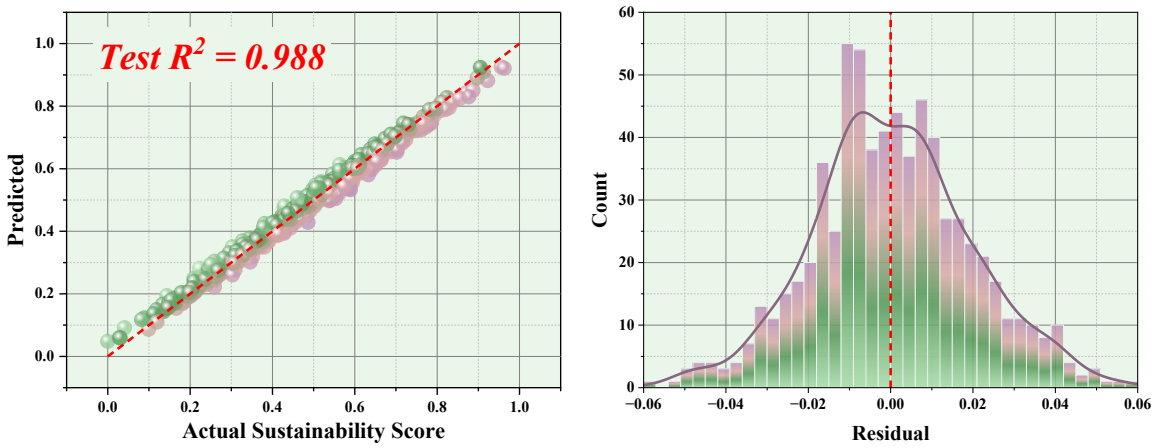


Fig.7. Accurate prediction of test sets and distribution of the residual errors.

4. Conclusion

In this research, the authors constructed an entirely data-driven multi-objective urban planning framework, which is capable of balancing sustainability, environmental quality, safety, and economic development. A Histogram Gradient Boosting (HistGB) surrogate model shows the non-linear interactions between 15 sustainability indicators. The sensitivity and ablation analysis indicated that green cover (ΔR^2 = 0.8196, Δ = 0.370) and the use of renewable energy (ΔR^2 = 0.3584, Δ = 0.244) are the strongest drivers of the sustainability of the urban areas with other indicators of urban sustainability having moderate or little influence. The biggest decrease in model performance (Test R^2 of 0.9877 to 0.1542) with the removal of environmental indicators confirms their importance. The Optimization Algorithms as the Novel Hybrid Kepler Optimization Algorithm (NOHKOA) was able to optimize the hyperparameters of the HistGB to have stable and precise surrogate predictions. The Dynamic Trade-off Coefficient Selection Multi-Objective (DTCSMO) algorithm was then used to optimize three conflicting goals, including sustainability maximization, the environmental load minimization, and the urban risk minimization, and process continuous and categorical land-use variables. The DTCSMO algorithm was able to produce 120

optimal solutions in pareto, giving varied and balanced trade-offs across the sustainability, environmental load and urban risk goals. The result of this smooth combination of surrogate modeling, hyperparameter optimization and multi-objective evolutionary optimization is a strong, explainable, and computationally efficient decision-support system in sustainable urban land-use planning. In general, the suggested framework can be useful to the urban planners and policy-makers in the focus of the significant role of the green infrastructure and integration of renewable energy sources. Its multi-objective, hybrid approach is adaptable and can be transferred to other cities experiencing intricate decisions in sustainable urban development and thus a viable and practical transferable methodology to evidence-based planning.

Declarations

Ethics Approval

Not applicable. This study uses a publicly available secondary dataset and does not involve human participants, animals, or identifiable personal data.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Data Availability

The dataset used in this study, *Sustainable Urban Planning & Landscape Dataset*, is publicly available on Kaggle at: https://www.kaggle.com/datasets/programmer3/sustainable-urban-planning-and-landscape-dataset

Competing Interests

The authors declare no competing interests.

Funding

The authors received no specific funding for this study.

Author Contributions

Cui Jing: Investigation, data curation, and methodology. Yiyun Wang: Methodology, software, and validation. Lina Zhao: Conceptualization, supervision, writing, review and editing. All authors read and approved the final manuscript.

Clinical Trial Number

Not applicable.

Nomenclature	
CV	cross-validation
DTCSMO	Dynamic Trade-off Coefficient Selection Multi-Objective
ES	Ecosystem Services
ER	Ecological Risk
IPCC	Intergovernmental Panel on Climate Change
HistGB	Histogram Gradient Boosting
NOHKOA	Novel Hybrid Kepler Optimization Algorithm
UNet	U-shaped Network

REFERENCES

[1] T. Sicuaio, P. Zhao, P. Pilesjo, A. Shindyapin, and A. Mansourian, "Sustainable and resilient land use planning: a multi-objective optimization approach," *ISPRS Int. J. Geoinf.*, vol. 13, no. 3, p. 99, 2024.

[2] C. Luan and R. Liu, "Integrated ecological and environmental spatial soft constraint mechanism: Urban multi-objective land use optimization under multiple scenarios," *J. Environ. Manage.*, vol. 385, p. 125661, 2025.

[3] Z. Wei, N. Du, and W. Yu, "Land use change and its driving factors in the ecological function area: A case study in the Hedong Region of the Gansu Province, China," *J. Arid Land*, vol. 16, no. 1, pp. 71–90, 2024.

[4] S. Xin *et al.*, "The contribution of multi-objective land use optimization to reducing ecological risk: A case study of the Lanzhou-Xining urban agglomeration," *Ecol. Indic.*, vol. 168, p. 112604, 2024.

[5] "Enhancing environmental sustainability in a critical region: climate change impacts on agriculture and tourism," *Civil Engineering Journal*, 2023.

[6] Y. Wang, X. Zhou, and K. Cao, “Low-carbon perspective-based multi-objective optimization of land use using improved NSGA-III in mega cities: A case study in Shenzhen,” *J. Environ. Manage.*, vol. 401, p. 128786, 2026.

[7] X. Chen *et al.*, “Land use optimization embedding in ecological suitability in the embryonic urban agglomeration,” *Land (Basel).*, vol. 12, no. 6, p. 1164, 2023.

[8] T. Pan, F. Su, F. Yan, V. Lyne, Z. Wang, and L. Xu, “Optimization of multi-objective multi-functional landuse zoning using a vector-based genetic algorithm,” *Cities*, vol. 137, p. 104256, 2023.

[9] Y. Dong, A. Subramanyam, and H. Wu, “Spatially explicit optimization of urban green infrastructure for multiple ecosystem services using deep learning surrogates,” *Environmental Modelling & Software*, p. 106758, 2025.

[10] P. T. Rodriguez-Gonzalez, A. Orozco-Calvillo, S. A. Tovar-Ortiz, E. Ruiz-Beltrán, and H. A. Olmos-Guerrero, “Multi-Objective Optimization of Socio-Ecological Systems for Global Warming Mitigation.,” *World (2673-4060)*, vol. 6, no. 4, 2025.

[11] Q. Cheng *et al.*, “Multi-objective spatial optimization by considering land use suitability in the Yangtze River Delta region,” *Journal of Geographical Sciences*, vol. 36, no. 1, pp. 45–78, 2026.

[12] S. Aosaf, M. A. Nayeem, A. Haque, and M. S. Rahman, “Computational Intelligence based Land-use Allocation Approaches for Mixed Use Areas,” *arXiv preprint arXiv:2508.15240*, 2025.

[13] Y. Wang, X. Zhou, and K. Cao, “Multi-objective land use optimization for enhanced urban livability using a NSGA-III-differential evolution algorithm,” *Land use policy*, vol. 161, p. 107880, 2026.

[14] R. Shu, Z. Wang, N. Guo, M. Wei, Y. Zou, and K. Hou, “Multi-scenario land use optimization simulation and ecosystem service value estimation based on fine-scale land survey data,” *Land (Basel).*, vol. 13, no. 4, p. 557, 2024.

[15] A. S. Cruz, L. R. Caldas, V. M. Mendes, J. C. Mendes, and L. E. G. Bastos, “Multi-objective optimization based on surrogate models for sustainable building design: A systematic literature review,” *Build. Environ.*, vol. 266, p. 112147, 2024.

[16] R. Shan, X. Jia, X. Su, Q. Xu, H. Ning, and J. Zhang, “Ai-driven multi-objective optimization and decision-making for urban building energy retrofit: advances, challenges, and systematic review,” *Applied Sciences*, vol. 15, no. 16, p. 8944, 2025.

[17] “Sustainable Urban Planning & Landscape Dataset”, Accessed: Feb. 23, 2026. [Online]. Available: https://www.kaggle.com/datasets/programmer3/sustainable-urban-planning-and-landscape-dataset?utm_source=chatgpt.com

[18] M. Saied, S. Guirguis, and M. Madbouly, “A comparative study of using boosting-based machine learning algorithms for IoT network intrusion detection,” *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, p. 177, 2023.

[19] R. Mohamed, M. Abdel-Basset, K. M. Sallam, I. M. Hezam, A. M. Alshamrani, and I. A. Hameed, “Novel hybrid kepler optimization algorithm for parameter estimation of photovoltaic modules,” *Sci. Rep.*, vol. 14, no. 1, p. 3453, 2024.

[20] Y. Hu, J. Ou, H. Wang, J. Zou, J. Zheng, and S. Yang, “Dynamic multiobjective optimization via an improved r-dominance relation and a novel prediction approach,” *Expert Syst. Appl.*, vol. 263, p. 125765, 2025.