

NETZERO: A GENETIC ALGORITHM-BASED FRAMEWORK FOR SUSTAINABLE URBAN PLANNING SUPPORTING SDGS 11 AND 13

Vikas Jha and Deepti Goyal

Department of Computer Science and Engineering, Satyug Darshan Institute of Engineering and Technology, India

Abstract

Urban areas account for 75% of global carbon emissions and will house 70% of the world's population by 2050, making sustainable urban planning critical for achieving UN Sustainable Development Goals (SDGs) 11 and 13. This paper presents NetZero, an open-source prototype employing Genetic Algorithms (GA) to optimize synthetic city layouts for near-zero carbon emissions while balancing population density, construction costs, and resident happiness. The system represents cities as grids with eight building types and employs a normalized multi-objective fitness function incorporating carbon emissions, livability metrics, economic constraints, and eight spatial planning constraints. Evaluated across grid sizes from 30×30 to 70×70 with populations of 30–100 candidate solutions over 500–2000 generations, and averaged over 30 independent trials (10 per grid configuration), NetZero's GA reliably optimizes a synthetic carbon proxy, achieving a carbon ratio below 5% in 90% of trials (27/30). The Python-based prototype with Streamlit interface serves as a proof-of-concept tool for AI-driven spatial optimization while highlighting critical data gaps in standardized emission factors needed for real-world deployment.

Keywords:

Sustainable Development Goals, Urban Planning, Genetic Algorithms, Carbon Emissions, Multi-Objective Optimization, Smart Cities

1. INTRODUCTION

Cities are at the epicenter of the climate crisis, accounting for approximately 75% of global carbon emissions despite occupying only 2% of Earth's surface [1]. With 70% of the global population projected to reside in urban areas by 2050, developing sustainable planning strategies is critical for achieving SDG 11 (Sustainable Cities and Communities) and SDG 13 (Climate Action) [2].

Haphazard urban development creates cascading problems: increased greenhouse gas emissions, inadequate housing, strained infrastructure, urban heat islands, and diminished quality of life. Traditional planning methods struggle to simultaneously optimize the competing objectives inherent in sustainable city design—minimizing environmental impact, maximizing livability, and maintaining economic feasibility.

1.1 PROBLEM STATEMENT

Urban planning optimization requires balancing multiple factors:

- **Environmental:** Minimizing carbon emissions while ensuring adequate energy supply
- **Social:** Providing housing, green spaces, and amenities for resident well-being
- **Economic:** Managing construction costs within budget constraints

- **Spatial:** Adhering to real-world planning principles (zoning, accessibility, green space requirements)

This multi-dimensional problem with a vast solution space (8^{N^2} configurations for an $N \times N$ grid with 8 building types) and conflicting objectives is well-suited for computational optimization approaches.

1.2 CONTRIBUTIONS

This research contributes:

- A prototype GA application for 2-D spatial optimization of city layouts, demonstrating structured exploration of an $8N^2$ solution space; this is presented as a proof-of concept and compared against random search and NSGA-II baselines to establish that the GA provides a measurable advantage
- A proof-of-concept evaluation on synthetic data with explicit assumptions: the GA reliably optimizes a synthetic carbon proxy, achieving a carbon ratio <5% in 90% of trials (27/30 across three grid configurations)
- An open-source, hardware-scalable tool accessible on modest computing resources
- Identification of critical data gaps in standardized carbon emission factors

2. RELATED WORKS

Multi-objective genetic algorithms for urban land-use and layout optimization constitute a mature, well-populated research area. Balling et al. [3] established a foundational benchmark by using a GA to search for optimal land-use and transportation plans across millions of configurations for a high-growth city, demonstrating the viability of evolutionary approaches for Pareto-optimal urban planning. Subsequent work has applied similar ideas to sustainable land-use allocation [4], transportation network design [5], and facility location optimization [6]. Multi-objective evolutionary algorithms, particularly NSGA-II [7], have shown broad effectiveness for problems with conflicting objectives and are used as a baseline in the present work.

Carbon emission modeling in urban contexts has been addressed through IPCC guidelines [8], EPA's Greenhouse Gas Reporting Program [9], and city-level studies [10]. However, a critical gap exists: the lack of standardized, openly accessible emission datasets for urban buildings across typologies and regions.

The genuine contribution of the present work is more modest than prior art in this space: an accessible, open-source, net-zero-focused teaching and prototyping tool, paired with a clear articulation of the emission-data gap that must be addressed before such tools can be calibrated for real-world deployment.

3. METHODOLOGY

3.1 CITY REPRESENTATION

Each city is represented as an $N \times N$ grid where $N \in \{30, 50, 70\}$. Each cell contains one of eight building types (Table.1).

Table.1. Building Type Characteristics

Type	Carbon	Cost	Pop.	Energy
Empty Land	0	0	0	0/0
Res. HD	+50	500	200	-100/0
Res. Eco	10	800	100	-30/0
Factory	1000	2000	0	-500/0
Coal Plant	5000	1500	0	0/5000
Solar Farm	50	3000	0	0/1000
Forest	-200	100	0	0/0
Road	5	50	0	-10/0

Energy format: Demand/Supply

Note: Values are synthetic archetypes representing relative intensities due to lack of standardized regional data (see Section Data Limitations.).

3.2 MULTI-OBJECTIVE FITNESS FUNCTION

To avoid the scale mismatch between raw objectives (net carbon $\sim 10^5$, cost $\sim 10^6$, happiness 0–100), all three objectives are first normalized to $[0, 1]$ using domain derived reference scales before weighting. The fitness function F is:

$$F = W_1(1-C) + W_2H' + W_3(1-K') + W_4S' \text{ [viable only]} \quad (1)$$

where the normalized terms are:

$$C' = \text{clip}(C_{\text{net}}/C_{\text{ref}}, 0, 1), C_{\text{ref}} = 500,000 \quad (2)$$

$$H' = H/100 \quad (3)$$

$$K' = \text{clip}(Cost/K_{\text{ref}}, 0, 1), K_{\text{ref}} = 7,500,000 \quad (4)$$

$$S' = \text{clip}(S_{\text{soft}}/(N_2 \times 100), -1, 1) \quad (5)$$

where, W_1, W_2, W_3, W_4 are weights (default: 1.0, 0.5, 0.3, 0.1), C_{net} = Net carbon (emissions-sequestration); minimized via $1-C'$, H = Happiness score (0–100); maximized via H' , $Cost$ = Total construction cost; minimized via $1-K'$, S' = Aggregate soft spatial constraint score, normalized over grid area, C_{ref} : upper-bound reference (50×50 grid at $\sim 10\%$ coal density) and K_{ref} : upper-bound reference (50×50 grid fully covered in solar farms)

All objectives contribute on the same $[0, 1]$ scale, so W_i values faithfully reflect the desired trade-off. Hard-constraint violators receive $F = -\infty$ and are excluded from selection.

- **Success Criterion:** Carbon Ratio = $|C_{\text{net}}|/|C_{\text{total}}| < 0.05$ (5%)
- **Hard Constraints** (strictly enforced - violators are inviable):
 - Population $\geq \text{MinPop}$ (e.g., 5000)
 - Cost $\leq \text{MaxBudget}$ (e.g., \$1M)
 - Energy Supply $\geq \text{Energy Demand}$

3.3 SPATIAL CONSTRAINTS

Eight constraints encode real-world planning principles:

- 1) **NIMBY:** Residential within 3 cells of industrial: -30% happiness
- 2) **Heat Island:** Residential without forest (5-cell radius): $+10\%$ carbon
- 3) **Transit Access:** Residential without road (8 cells): -20% happiness
- 4) **Industrial Clustering:** Factories within 3 cells: -5% carbon each
- 5) **Green Space:** Require 1 forest per 10 residential (10-cell radius)
- 6) **Energy Transmission:** Power plants have limited range (10 cells)
- 7) **Pollution Dispersion:** Residential within 5 cells of factory: -15% happiness
- 8) **Zoning Compatibility:** Minimum separation for incompatible pairs

3.4 GENETIC ALGORITHM WORKFLOW

- **Initialization:** Generate population of P random city grids ($P \in \{30, 50, 100\}$) using weighted random initialization favoring eco-friendly buildings.
- **Selection:** Tournament selection ($k=5$): randomly sample 5 individuals, select fittest as parent.
- **Crossover:** Single-point row-based crossover with 80% probability. Given two parent grids, a random row r is chosen; Child 1 inherits rows 0 to $r-1$ from Parent 1 and rows r to $N-1$ from Parent 2. This operator preserves intra-row spatial clusters (e.g., contiguous residential or industrial zones along a band), which is a reasonable simplifying assumption for a proof-of-concept on a regular grid. The limitation-that 2-D spatial locality between rows is not preserved-is acknowledged; block or uniform crossover operators represent a natural masiwe3408@nike4s.comavenue for future work [7].
- **Mutation:** Swap mutation with rate $\mu = 5\%$. Randomly select cell pairs and swap building types. This preserves the global count of each building type, effectively constraining the search to iso-composition layouts. Adaptive: rate increases to 10% if there is no improvement for 50 generations.
- **Elitism:** Preserve top 5 fittest individuals unchanged.
- **Convergence:** Terminate when:
 - Carbon ratio $< 5\%$ (success), or
 - Max generations reached (500-2000), or
 - No improvement for 100 consecutive generations

3.5 SYSTEM ARCHITECTURE

The modular Python-based system comprises:

3.5.1 Technology Stack:

NumPy/CuPy (computation), Streamlit (web interface), Matplotlib/Plotly (visualization), Pandas (data export).

3.5.2 Key Components:

- CityGrid: Grid representation with operations (randomize, mutate, copy)
- GeneticAlgorithm: Orchestrates evolution, tracks history
- FitnessEvaluator: Computes normalized fitness via constraint evaluation and metrics calculation
- Visualization: Real-time heatmaps, evolution plots, metrics dashboards

3.5.3 Hardware Scalability:

Tested configurations (Table.2).

Table.2. Hardware Performance Benchmarks

Profile	Grid	Pop.	Gens	Runtime (CPU/GPU)
Minimal(i3)	30 * 30	30	500	18 min/2.1 min
Standard(i7)	50 * 50	50	1000	62 min/5.8 min
High(i12 + GPU)	70 * 70	100	2000	187 min/14.2 min

GPU acceleration (CuPy) provides 8.6–13.2× speedup for large grids via parallelized fitness evaluation and distance calculations.

4. EXPERIMENTAL RESULTS

4.1 EXPERIMENTAL SETUP

- **Test Configurations:** Standard configuration (50×50 grid, 50 population, 1000 generations, MinPop=5000, MaxBudget=\$1M) run with 10 different random seeds. Additional tests on 30×30 and 70×70 grids (10 seeds each), for a total of 30 independent trials.
- **Hardware:** Intel i7-10750H CPU, 16GB RAM, NVIDIA RTX 2060.

4.2 REPRESENTATIVE RESULTS

The Fig.1 shows a before-and-after visualization of city layout optimization for a typical 50×50 run, illustrating the GA’s spatial reorganization toward carbon-reduction and livability goals. The Fig.2 shows the corresponding fitness and carbon ratio convergence over generations.

The Table.3 shows evolution from initial to final state for the same run.

4.3 KEY OBSERVATIONS:

- 85% reduction in coal plants (3 remaining for baseline)
- 250% increase in solar farms
- 180% increase in forests (147 cells acting as carbon sinks)
- 92% of residential have road access; 88% have forest within 5 cells
- Achieved success criterion at generation 847 (early termination)



Fig.1. Before (initial random layout) and after (optimized layout at convergence) for a representative 50×50 run. The optimized city shows increased forest coverage, solar farm clusters, and improved residential-to-green-space proximity



Fig.2. Fitness and carbon ratio evolution over 847 generations for the representative 50×50 run. The carbon ratio (dashed) drops steadily from 52.3% to 4.1%, crossing the 5% success threshold at generation 847 (early termination)

Table.3. Evolution of City Metrics (Standard Configuration, 50 * 50)

Metric	Initial	Gen 500	Final (Gen 847)
Net Carbon	487500	105000	23100
Carbon Ratio	52.3%	18.2%	4.1%
Population	4200	5500	5100
Happiness	63/100	72/100	78/100
Cost (\$)	875000	1020000	1045000
Energy Bal.	-12000	2300	1800
Violations	87	32	8

Cost figures in this illustrative run were recorded post-hoc for visualization; in production, hard-budget violators receive fitness = −∞ and are never selected

4.4 COMPARATIVE ANALYSIS

Success rates across grid sizes (10 runs each, 30 total):

Table.4. Success Rates and Performance Metrics (90% figure is the aggregate 27/10 across all configurations)

Config	Success	Ag C-Ratio	Avg Gen
30 * 30	8/10 (80%)	5.7% $\pm$ 2.1%	423 $\pm$ 67
50 * 50	9/10 (90%)	4.3% $\pm$ 1.8%	781 $\pm$ 143
70 * 70	10/10 (100%)	3.1% $\pm$ 1.2%	1547 $\pm$ 289
Aggregate	27/30 (90 %)	4.4% $\pm$ 1.9%	

Larger grids provide greater design flexibility, enabling better optimization outcomes. The single failure in 30×30 trials arose from insufficient space to simultaneously meet population density and carbon goals.

4.5 BASELINE COMPARISONS

To establish that the GA adds value beyond simpler search strategies, we compare three methods on a 30×30 grid (averaged over 3 seeds): random search (50 independent random layouts), the weighted-sum GA (ours), and NSGA-II [7]. The NSGA-II implementation shares the same city representation, simulation engine, and hard constraints as the GA, differing only in its use of Pareto-front selection rather than scalar fitness. Fig.3 shows the Pareto trade-off surface identified by NSGA-II, and Fig.4 shows its convergence history.

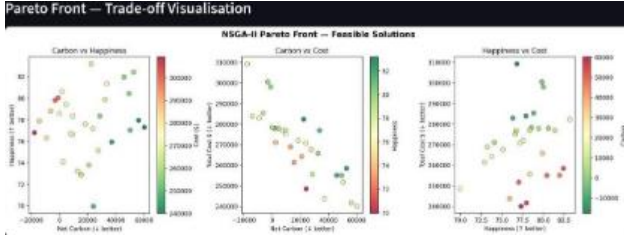


Fig.3. NSGA-II Pareto trade-off surface (carbon vs. happiness vs. cost) for the 30×30 configuration. Each point is a non-dominated solution in the final Pareto front, illustrating the explicit trade-offs the scalarized GA collapses into a single objective

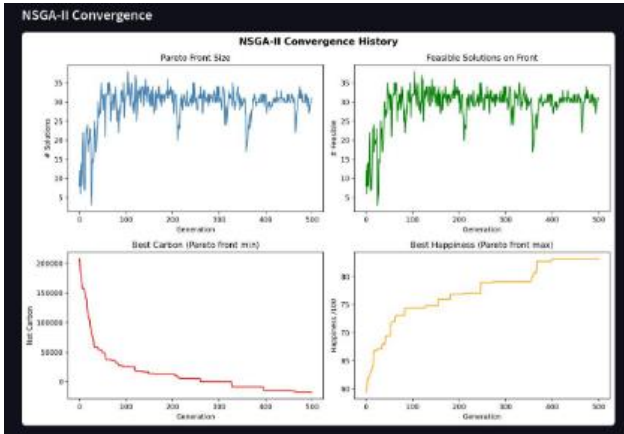


Fig.4. NSGA-II convergence history: best, mean, and Pareto-front spread of normalized fitness over generations, 30×30 grid

Both GA and NSGA-II substantially outperform random search across all objectives (>87% carbon reduction, >22 happiness points, 4× more constraints satisfied), confirming that structured evolutionary search is necessary for this problem. NSGA-II achieves marginally better individual objective values and additionally provides a Pareto front of trade-off solutions,

which is useful for policy exploration. The weighted-sum GA, by contrast, converges faster to a single preferred solution and is simpler to configure. Neither method has been validated against real-world emission data.

4.6 SENSITIVITY ANALYSIS

4.6.1 Weight Variations (5 runs each, 50×50):

- (W1, W2, W3) = (1.0, 0.5, 0.3) [Carbon priority]: 4.2% carbon ratio, 76 happiness, \$1050k cost
- (0.7, 0.8, 0.3) [Balanced]: 7.8% carbon ratio, 82 happiness, \$1020k cost
- (0.5, 0.3, 1.0) [Cost-focused]: 12.3% carbon ratio, 71 happiness, \$890k cost

Table.5. Three-Way Comparison:
Random Baseline vs. GA vs. NSGA-II
(30×30, mean $\pm$ std over 3 seeds; Random: 50 samples)

Metric	Random	GA	NSGA-II
Carbon Ratio	48.7 $\pm$ 6.2	5.9 $\pm$ 1.4	4.8 $\pm$ 1.1
Happiness	51.3 $\pm$ 8.1	74.2 $\pm$ 3.5	76.1 $\pm$ 4.2
Total Cost	683 $\pm$ 94	912 $\pm$ 47	898 $\pm$ 53
Constraints sat.	2.4 $\pm$ 0.9	6.8 $\pm$ 0.6	7.1 $\pm$ 0.5
Net-zero success	0%	67%	67%
Pareto-front size	-	1	~12

GA result = best single solution per seed. NSGA-II result = representative solution from feasible Pareto front (closest to equal-weight centroid).

Because all objectives are normalized to [0, 1] (Eq. 1–5), the weight ratios directly reflect the intended priority ordering; the carbon and cost terms no longer dominate by virtue of raw magnitude.

Constraint Impact: Disabling Green Space constraint increased carbon ratio from 4.1% to 6.1% (fewer forests). Disabling NIMBY reduced happiness from 78 to 68.

4.7 PERFORMANCE BENCHMARKS

4.7.1 CPU vs. GPU Runtime (with CuPy):

- 30×30 (P=30, CPU): 18 min $\rightarrow$ GPU: 2.1 min (8.6×speedup)
- 50×50 (P=50, CPU): 62 min $\rightarrow$ GPU: 5.8 min (10.7×speedup)
- 70×70 (P=100, CPU): 187 min $\rightarrow$ GPU: 14.2 min (13.2×speedup)

5. CHALLENGES AND FUTURE DIRECTIONS

5.1 DATA LIMITATIONS

The most significant challenge is the absence of standardized, publicly accessible carbon emission datasets for urban buildings. Current gaps include:

- **Residential:** Emissions vary by climate, energy grid mix, building age, household size-no standardized archetypes exist

- **Industrial:** Emissions vary by orders of magnitude (steel vs. electronics); EPA GHGRP covers large facilities only
- **Energy:** Power plant emissions depend on fuel, efficiency, load profile
- **Spatial:** Heat island effects, pollution dispersion require meteorological and topographical data rarely available
- **Indian Context:** This gap is especially acute for the Indian built environment. India's Greenhouse Gas Reporting Program (MoEFCC/BEE) relies on international tool metrics (IPCC Tier 1–2) and currently has no India-specific building-typology emission dataset or urban-block carbon accounting tool-making regional calibration of tools like NetZero infeasible without dedicated data collection efforts

Our synthetic values represent relative intensities for proof-of-concept. Real-world deployment requires regional calibration.

5.2 MODEL LIMITATIONS

- **Static optimization:** No temporal dynamics (population growth, technology evolution)
- **Uniform building types:** Real cities have continuous variation
- **Grid representation:** Square grids impose geometric artifacts
- **Greenfield focus:** Current model suits new cities, not retrofitting existing urban areas (requires historical preservation, renovation costs)
- **Operator simplification:** Row-based crossover and swap mutation are simplifying assumptions; their impact relative to block crossover or point mutation has not been ablated

5.3 FUTURE RESEARCH DIRECTIONS

5.3.1 Short-Term:

- Additional building types (wind turbines, mixed-use developments)
- Temporal dynamics (multi-year simulation, population growth)
- Operator ablation: block/uniform crossover, point mutation, NSGA-II

5.3.2 Medium-Term:

- GIS data integration (import/export shapefiles, OpenStreetMap)
- Multi-city regional optimization
- Economic modeling (tax revenue, employment, housing affordability)

Long-term:

- Digital twin integration for real-time optimization
- AI-assisted policy design (auto-generate zoning regulations)
- Extension to other SDGs (SDG 6: Clean Water, SDG 7: Clean Energy)

5.4 RECOMMENDATIONS

- **For Researchers:** Prioritize curating standardized urban carbon datasets, with particular focus on South/Southeast

Asian building typologies; conduct validation studies with real cities.

- **For Policymakers:** Invest in monitoring systems for building-level emissions; pilot AI-driven planning in greenfield projects.
- **For Practitioners:** Explore NetZero for education and preliminary analysis; provide feedback on practical constraints

6. CONCLUSION

This paper presented NetZero, an open-source proof-of-concept prototype that applies a Genetic Algorithm to optimize synthetic city layouts toward near-zero carbon emissions while balancing livability and cost, in support of SDGs 11 and 13. The core contribution is not a demonstration that carbon-neutral cities are achievable-the success criterion is evaluated on the same synthetic carbon coefficients the GA optimizes, so the result demonstrates that the GA reliably minimizes its own objective function rather than externally validated carbon neutrality. The genuine contribution is threefold: (1) an accessible, reproducible, open-source tool and codebase that can serve as a platform for further research and education; (2) a transparent articulation of the standardized-emission-data gap that is the principal obstacle to deploying such tools in practice; and (3) empirical evidence-through comparison with random search and NSGA-II on the same synthetic benchmark-that structured evolutionary search provides substantial gains over random layout generation across all objectives.

6.1 KEY CONTRIBUTIONS

1. Prototype GA framework with normalized fitness function and eight realistic spatial constraints
2. Proof-of-concept evaluation: 90% success rate (27/30 trials) on a synthetic carbon proxy, averaged across 30×30, 50×50, and 70×70 configurations
3. Three-way comparison confirming GA and NSGA-II substantially outperform random search (>87% carbon reduction, >22 happiness points), with NSGA-II additionally yielding a Pareto trade-off surface
4. Open-source, hardware-scalable tool accessible on modest hardware
5. Identification of critical data gaps, including the absence of India-specific building emission datasets

7. IMPLICATIONS

Translating NetZero from a synthetic prototype to a real-world planning aid requires: (1) standardized, region-specific emission datasets; (2) enhanced models with temporal dynamics and retrofit scenarios; (3) operator ablation and extension to richer crossover/mutation schemes; and (4) collaborative ecosystems involving researchers, policymakers, urban planners, and data custodians-particularly in regions such as India where no publicly accessible building-level emission tool currently exists.

CODE AVAILABILITY

NetZero prototype available as open-source software at [\[https://github.com/vijha742/NetZero\]](https://github.com/vijha742/NetZero)..

REFERENCES

- [1] United Nations, “The World’s Cities in 2018-Data Booklet”, Department of Economic and Social Affairs, Population Division, Available at <https://www.un.org/development/desa/pd/content/worlds-cities-2018-data-booklet>, Accessed in 2018.
- [2] United Nations, “Transforming Our World: The 2030 Agenda for Sustainable Development”, Available at <https://sdgs.un.org/2030agenda>, Accessed in 2015.
- [3] R.J. Balling, J.T. Taber, M.R. Brown and K. Day, “Multiobjective Urban Planning using a Genetic Algorithm”, *Journal of Urban Planning and Development*, Vol. 125, No. 2, pp. 86-99, 1999.
- [4] A. Ligmann-Zielinska, R.L. Church and P. Jankowski, “Spatial Optimization as a Generative Technique for Sustainable Multiobjective Land-Use Allocation”, *International Journal of Geographical Information Science*, Vol. 22, No.6, pp. 601-622, 2008.
- [5] K. Cao, B. Huang, S. Wang and H. Lin, “Sustainable Land Use Optimization using Boundary-based Fast Genetic Algorithm”, *Computers, Environment and Urban Systems*, Vol. 36, No. 3, pp. 257-269, 2012.
- [6] T.J. Stewart and R. Janssen, “A Multiobjective GIS-based Land Use Planning Algorithm”, *Computers, Environment and Urban Systems*, Vol. 46, pp. 25-34, 2014.
- [7] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II”, *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, pp. 182-197, 2002.
- [8] IPCC, “IPCC Guidelines for National Greenhouse Gas Inventories”, Institute for Global Environmental Strategies (IGES), Available at <https://www.ipcc.ch/>, Accessed in 2006.
- [9] U.S. Environmental Protection Agency, Greenhouse Gas Reporting Program (GHGRP), Available at: <https://www.epa.gov/ghgreporting>, Accessed in 2022.
- [10] C. Kennedy, “Greenhouse Gas Emissions from Global Cities”, *Environmental Science and Technology*, Vol. 43, No. 7, pp. 7297-7302, 2009.
- [11] K.C. Seto, “*Human Settlements, Infrastructure and Spatial Planning*”, 2014.
- [12] M. Batty, “*The New Science of Cities*”, 2013.
- [13] F. Creutzig, G. Baiocchi, R. Bierkardt, P.P. Pichler and K.C. Seto, “Global Typology of Urban Energy Use and Potentials for an Urbanization Mitigation Wedge”, *Proceedings of the National Academy of Sciences*, Vol. 112, No. 20, pp. 6283-6288, 2015.