ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

Dynamic Equity-Aware Pareto Optimization for Predictive Resource Allocation in Smart City Infrastructure

Xinyi Pan

University of Toronto, Toronto, ON, Canada

Abstract: *Static resource-allocation models can overlook spatial inequity and changing urban demand. We propose a closed-loop, online spatiotemporal multi-objective framework that jointly addresses equity and efficiency in smart city infrastructure. A CNN-BiGRU forecaster with additive attention estimates demand means and variances; a self-adaptive estimation-of-distribution evolutionary optimizer searches for allocation policies along the Pareto frontier of spatial accessibility inequality and service coverage minus cost. A sliding-window Kolmogorov-Smirnov detector identifies distribution shifts and triggers re-optimization only in affected subregions. Finally, a decision-support layer selects a policy from the Pareto set using an equity-adjusted social welfare function combining an Atkinson-based term and a Gini penalty. Predictions, allocation decisions, and observed service outcomes form a feedback loop. Evaluation on synthetic demand shifts and New York City taxi records, as reported in this study, compares forecasting, solution quality, equity, and computation time with established baselines. The framework offers an integrated approach to uncertainty-aware forecasting, adaptive resource allocation, and transparent equity-efficiency trade-offs.*

Keywords: *Smart City Infrastructure, Urban Resource Allocation, Multi-Objective Optimization, Pareto Optimization, Spatial Equity, Spatiotemporal Forecasting, Distribution Shift*

1. Introduction

Rapid urbanization strains emergency services, transit, waste management, and energy systems. Smart cities use digital sensing and computation to improve service delivery [1], yet benefits can be distributed unevenly across communities [2]. Public services also respond to non-stationary demand shaped by seasonal activity, neighborhood conditions, and infrastructure changes. Static, single-objective allocation models may overlook these spatial and temporal patterns [3] and distribute resources without an explicit account of spatial equity.

Urban service needs vary by hour, season, and neighborhood and can exhibit spatial autocorrelation [4]. Maximizing aggregate coverage may concentrate resources in dense areas, while improving access in underserved areas can increase delivery costs. Conversely, a policy chosen for fairness alone may not reflect the expense or practical capacity of service provision. Allocation therefore requires adaptive decisions across competing equity and efficiency objectives, especially where patterns of need shift faster than fixed planning cycles.

We propose a closed-loop framework within an Urban Data Platform's Analytics Engine. It combines uncertainty-aware CNN-BiGRU demand forecasts, EDA-based evolutionary Pareto optimization, Kolmogorov-Smirnov shift detection with subregional re-optimization, and an equity-adjusted social welfare policy selector. Equity is measured through the spatial Gini coefficient of accessibility; efficiency reflects coverage minus delivery costs. Streaming data and service outcomes continually update the allocation process.

The contributions are (1) integrating probabilistic spatiotemporal forecasts with multi-objective allocation; (2) adapting to non-stationary demand through selective rather than global re-optimization; and (3) providing an adjustable welfare-based decision layer that makes the equity-efficiency trade-off explicit. Sections 2-3 review prior work and definitions; Sections 4-6 present the framework, evaluation, and results; Sections 7-8 discuss limitations and conclude.

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

2. Related Work

Three research streams inform this framework: spatiotemporal forecasting, multi-objective allocation, and spatial equity.

2.1 Spatiotemporal Demand Forecasting

ARIMA-type forecasting can miss nonlinear urban dynamics. CNNs represent local spatial patterns [5], while recurrent networks model temporal dependencies [4]; hybrid architectures combine both [6,7]. Most cited models emphasize point forecasts. Our CNN-BiGRU forecaster also estimates predictive variance so downstream allocations can reflect uncertainty.

2.2 Multi-Objective Optimization for Resource Allocation

Resource allocation problems in urban systems are inherently multi-objective, often requiring a trade-off between efficiency (e.g., maximizing coverage or minimizing cost) and equity (e.g., guaranteeing fair access across neighborhoods). Conventional approaches have employed weighted-sum methods or ϵ-constraint methods to scalarize the multi-objective problem, but these necessitate prior knowledge of the decision-maker's preferences and may overlook non-convex areas of the Pareto frontier. Evolutionary multi-objective optimization (EMO) algorithms, such as NSGA-II [8] and MOEA/D [9], have been widely applied to discover the Pareto frontier without such assumptions. For instance, a multi-objective optimization algorithm has been applied to dynamic resource allocation in cloud computing [3], and a multi-objective safe reinforcement learning method was proposed for equity-preserving public health resource allocation [10]. However, these methods typically assume a stationary environment and require a full recomputation when the demand distribution shifts. Our framework introduces a self-adaptive evolutionary optimizer with an estimation-of-distribution (EDA) search operator, which is particularly well-suited for online optimization because it builds a probabilistic model of promising solutions and samples new candidates efficiently. The EDA approach has been applied to multi-objective ship route optimization in non-stationary tidal environments [11], but its application to urban resource allocation is without precedent.

2.3 Equity-Aware Decision-Making in Smart Cities

Smart-city research documents unequal access to digital and public infrastructure [1,2]. Equity-oriented optimization has been developed for urban water resources [12] and dynamically constrained resource allocation [13]. Our approach combines a spatial accessibility metric with temporally adaptive optimization and an adjustable welfare-based choice among Pareto policies.

2.4 Comparison with Existing Works

The framework's proposed contribution is the integration of probabilistic forecasts, EDA-based Pareto search, selective updates under distribution shift, and a transparent equity-adjusted policy selector in one feedback loop.

3. Preliminaries and Equity Metrics

We define demand uncertainty, Pareto dominance, and the spatial equity measures used below.

3.1 Spatiotemporal Non-Stationarity and Probabilistic Forecasting Foundations

Urban demand data present complex spatiotemporal dependencies, which challenge conventional forecasting methods. The demand for a service at a specific location and time, denoted as $Y_{s,t}$, is influenced by both its historical values at the same location and the contemporaneous values at neighboring locations. This spatial autocorrelation can be formally expressed as:

$$\mathrm{Cov}\left(Y_{s,t}, Y_{s',t}\right) = f\left(d(s,s')\right) \tag{1}$$

where $d(s,s')$ represents the distance between spatial locations s and s', and f is a decreasing function of distance, reflecting Tobler's first law of geography that "everything is related to everything else, but near things are more related than distant things."

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

Urban demand is non-stationary under seasonality, events, infrastructure changes, and policy interventions. We therefore model a conditional predictive distribution rather than a point estimate:

$$Y_{s,t} \sim P\left(y \middle| \mathcal{H}_{s,t}\right) \tag{2}$$

where $\mathcal{H}_{s,t}$ denotes the historical information available up to time t for location s and its spatial neighbors. This probabilistic formulation permits the forecaster to quantify uncertainty, typically by generating both a mean $\hat{y}_{t+1}$ and a variance $\hat{\sigma}_{t+1}^2$ for each spatiotemporal cell. A common representation of the prediction interval is:

$$\hat{y}_{t+1} \pm z_{\alpha/2} \cdot \hat{\sigma}_{t+1} \tag{3}$$

where $z_{\alpha/2}$ is the critical value from the standard normal distribution corresponding to a desired confidence level $(1-\alpha)$. This uncertainty quantification is critical for subsequent optimization, as it permits the system to make risk-aware allocation decisions instead of relying on deterministic forecasts.

3.2 Principles of Multi-Objective Optimization and Pareto Dominance

Allocation decisions can trade off coverage, cost, and equity. A general multi-objective problem is:

$$\min_{\mathbf{x}} \mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})]^{\mathsf{T}} \tag{4}$$

where $\mathbf{x} \in \Omega$ is a decision vector representing an allocation policy, and $f_i(\mathbf{x})$ is the i-th objective function to be minimized. In contrast to single-objective optimization, there is generally no solitary solution that minimizes all objectives at the same time. Instead, we seek the set of Pareto-optimal solutions, defined through the concept of Pareto dominance.

A solution **a** is said to dominate another solution **b** (denoted $\mathbf{a} \preccurlyeq \mathbf{b}$) if it is at least as good in all objectives and strictly better in at least one:

$$\mathbf{a} \preccurlyeq \mathbf{b} \Leftrightarrow \forall i: f_i(\mathbf{a}) \leq f_i(\mathbf{b}) \wedge \exists j: f_j(\mathbf{a}) < f_j(\mathbf{b}) \tag{5}$$

All non-dominated solutions collectively constitute the Pareto frontier.

$$\mathcal{P}^* = \{\mathbf{x} \in \Omega \mid \nexists \mathbf{x}' \in \Omega: \mathbf{x}' \preccurlyeq \mathbf{x}\} \tag{6}$$

NSGA-II [8] and MOEA/D [9] maintain populations of candidate solutions to explore the Pareto frontier; online allocation requires adaptation when objectives shift with demand.

3.3 Quantitative Metrics for Spatial Equity and Social Welfare

The spatial Gini coefficient quantifies inequality across accessibility values:

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2n^2 \bar{x}} \tag{7}$$

where x_i is the accessibility measure (e.g., service coverage or resource allocation) for spatial unit i, $\bar{x}$ is the mean accessibility across all n units, and G ranges from 0 (perfect equality) to 1 (maximum inequality). A lower Gini coefficient indicates a more equitable distribution of resources.

An inequality-averse social welfare term incorporates the parameter ρ:

$$W = \left(\frac{1}{n}\sum_{i=1}^{n} u_i^{1-\rho}\right)^{\frac{1}{1-\rho}} \tag{8}$$

where u_i is the utility (e.g., accessibility) of spatial unit i, and $\rho \geq 0$ is the inequality aversion parameter. When ρ = 0, the welfare function is equivalent to the arithmetic mean; as ρ rises, the function becomes more responsive to shifts from the poor to the rich, thereby imposing a stricter penalty on inequality.

Accessibility is related to service supply and demand; the study uses the following measure:

$$A_s = S_s \times D_s \tag{9}$$

where S_s represents the service supply (e.g., number of ambulances or service capacity) allocated to location s, and D_s represents the demand at that location. This multiplicative formulation reflects the

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

intuitive concept that accessibility is elevated when supply and demand are aligned, and diminished when they are not. Together, these metrics supply the theoretical foundation for quantifying the equity objective in our optimization framework, which consequently permits the system to explicitly trade off efficiency against fairness.

4. Proposed Framework

Figure 1 locates four modules in the Urban Data Platform's Analytics Engine: probabilistic demand forecasting, self-adaptive EDA optimization, shift detection with local updates, and welfare-based decision support. Data ingestion and service delivery close the loop.

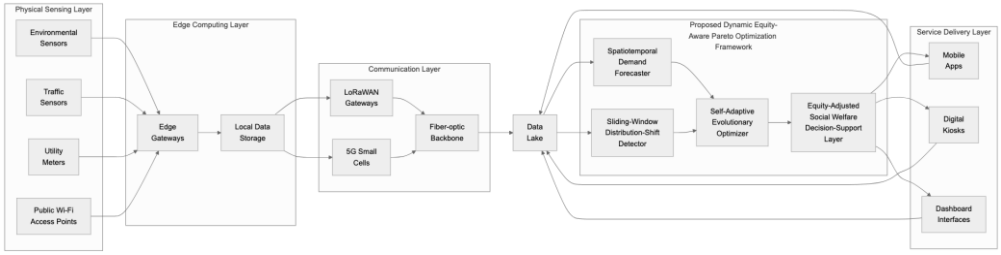


Figure 1: System Architecture of Equity-Aware Smart City Infrastructure.

The framework operates at discrete time steps $t = 1,2, \dots, T$, where each time step corresponds to a decision epoch (e.g., one hour or one day). At each time step, the system ingests streaming urban data, produces a forecast of future demand, optimizes the allocation policy, detects any distribution shifts, and dispatches the selected policy to the service delivery layer. The actual outcomes are subsequently returned to the data lake, which thereby completes the feedback loop and supports ongoing adaptation.

4.1 Spatiotemporal Demand Forecaster with Uncertainty Quantification

The initial module constitutes a probabilistic spatiotemporal predictor that estimates both the average and the variance of forthcoming service demand at each spatial cell. We denote the spatial domain as a grid of S cells, indexed by $s = 1,2, \dots, S$, where each cell corresponds to a census tract or a fixed-size grid cell. The historical demand observations up to time t are captured by a tensor $\mathbf{Y}_{1:t} \in \mathbb{R}^{S \times T_{\text{hist}}}$, where T_{hist} denotes the length of the historical window.

The forecaster architecture integrates a convolutional neural network (CNN) for spatial feature extraction, a bidirectional gated recurrent unit (BiGRU) for temporal sequence modeling, and an additive attention mechanism that emphasizes relevant historical time steps. The CNN module applies a set of 2D convolutional filters to the spatial grid, thereby capturing local spatial patterns, including demand hotspots and spatial autocorrelation. The CNN yields a spatial feature map; d_{spatial} denotes the number of spatial features per cell.

The BiGRU component processes the temporal sequence of spatial feature maps. For each spatial cell s, the BiGRU processes the sequence of feature vectors $\mathbf{h}_{s,1}, \mathbf{h}_{s,2}, \dots, \mathbf{h}_{s,T_{\text{hist}}}$ in both forward and backward directions, thereby generating a hidden state $\mathbf{h}_{s,t}^{\text{BiGRU}}$ that encodes temporal dependencies from both past and future contexts. The additive attention mechanism then computes a context vector $\mathbf{c}_{s,t}$ by weighting the hidden states according to their relevance to the current prediction:

$$\mathbf{c}_{s,t} = \sum_{\tau=1}^{T_{\text{hist}}} \alpha_{s,t,\tau}\, \mathbf{h}_{s,\tau}^{\text{BiGRU}} \tag{10}$$

where the attention weights $\alpha_{s,t,\tau}$ are computed as:

$$\alpha_{s,t,\tau} = \frac{\exp(\mathbf{v}^{\top}\tanh(\mathbf{W}_1\mathbf{h}_{s,\tau}^{\text{BiGRU}} + \mathbf{W}_2\mathbf{h}_{s,t}^{\text{BiGRU}}))}{\sum_{\tau'=1}^{T_{\text{hist}}} \exp\left(\mathbf{v}^{\top}\tanh(\mathbf{W}_1\mathbf{h}_{s,\tau'}^{\text{BiGRU}} + \mathbf{W}_2\mathbf{h}_{s,t}^{\text{BiGRU}})\right)} \tag{11}$$

Here, $\mathbf{v}$, $\mathbf{W}_1$, and $\mathbf{W}_2$ are adjustable parameters. The context vector $\mathbf{c}_{s,t}$ is then concatenated with the current hidden state and passed through a fully connected layer to produce the forecast parameters.

The forecaster outputs a Gaussian distribution for each spatial cell, parameterized by a mean $\hat{\mu}_{s,t+1}$

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

and a variance $\hat{\sigma}_{s,t+1}^2$:

$$\hat{\mu}_{s,t+1}, \hat{\sigma}_{s,t+1}^2 = f_{\text{forecast}}(\mathbf{Y}_{1:t}, \mathbf{X}_{1:t}) \tag{12}$$

where $\mathbf{X}_{1:t}$ represents exogenous features such as time-of-day, day-of-week, weather conditions, and socioeconomic indicators. The model is trained by minimizing the negative log-likelihood loss:

$$\mathcal{L} = \sum_{s=1}^{S}\sum_{t=1}^{T}\left[\frac{1}{2}\log(2\pi\hat{\sigma}_{s,t}^2) + \frac{(Y_{s,t} - \hat{\mu}_{s,t})^2}{2\hat{\sigma}_{s,t}^2}\right] \tag{13}$$

Forecast uncertainty is passed to the allocation optimizer.

4.2 Self-Adaptive Evolutionary Optimizer with EDA Search

The second module is a multi-objective evolutionary optimizer that develops a population of allocation policies toward the equity-efficiency Pareto frontier. An allocation policy is defined by a vector $\mathbf{p} \in [0,1]^S$, where p_s indicates the proportion of accessible resources assigned to spatial cell s. The total resources are scaled so that $\sum_{s=1}^{S} p_s = 1$.

The optimizer simultaneously optimizes two objectives. The first objective is efficiency, measured as the expected aggregate service coverage minus delivery costs:

$$f_{\text{eff}}(\mathbf{p}) = \sum_{s=1}^{S}\left[\min\left(\hat{\mu}_{s,t+1} \cdot p_s, \hat{\mu}_{s,t+1}\right) - c_s \cdot p_s\right] \tag{14}$$

where $\hat{\mu}_{s,t+1}$ is the forecasted mean demand, c_s is the unit cost of delivering service to cell s, and the min operator ensures that coverage does not exceed demand. The second objective is equity, measured as the negative of the spatial Gini coefficient of expected accessibility:

$$f_{\text{eq}}(\mathbf{p}) = -G\left(\mathbf{a}(\mathbf{p})\right) \tag{15}$$

where $\mathbf{a}(\mathbf{p}) = [a_1(\mathbf{p}), a_2(\mathbf{p}), \ldots, a_S(\mathbf{p})]^\top$ is the vector of expected accessibility values, with $a_s(\mathbf{p}) = \hat{\mu}_{s,t+1} \cdot p_s$, and $G(\cdot)$ is the Gini coefficient defined in Equation 7. The negative sign converts the minimization of inequality into a maximization objective, consistent with the Pareto dominance framework.

The optimizer keeps a set of N candidate policies. At each generation g, the population is evaluated on both objectives, and non-dominated sorting is applied to identify the Pareto frontier. The central novelty is the adoption of a self-adaptive estimation-of-distribution (EDA) search operator to generate offspring. Rather than employing conventional crossover and mutation operators, the EDA constructs a probabilistic model of the best-performing solutions and draws new candidates from this model.

Specifically, at generation g, we select the top 20% of the population based on non-dominated rank and crowding distance. These elite solutions are applied to fit a multivariate Gaussian mixture model (GMM) with K_g components.

$$\hat{\mathcal{P}}_{\text{elite}} = \left\{\mathbf{p}^{(i)} \mid \text{rank}\left(\mathbf{p}^{(i)}\right) \leq 0.2N\right\} \tag{16}$$

$$\mathcal{M}_g = \sum_{k=1}^{K_g} \pi_k \, \mathcal{N}(\mathbf{m}_k, \Sigma_k) \tag{17}$$

where π_k, $\mathbf{m}_k$, and Σ_k are the mixing coefficient, mean vector, and covariance matrix of the k-th Gaussian component, estimated from the elite set using the expectation-maximization algorithm. New candidate policies are then sampled from this GMM:

$$\mathbf{p}_{\text{new}} \sim \mathcal{M}_g \tag{18}$$

The quantity of GMM components K_g and the mutation rate η_g are tuned online according to the hypervolume gain of the Pareto front observed during the past $L = 5$ generations. Let HV_g denote the hypervolume of the Pareto frontier at generation g, computed with respect to a reference point $\mathbf{r} = [0, -1]^\top$. The hypervolume improvement is:

$$\Delta\text{HV}_g = \text{HV}_g - \text{HV}_{g-L} \tag{19}$$

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

When $\Delta HV_g < \delta = 0.01$, the mutation rate increases by 0.05 (maximum 0.3) and the GMM component count decreases by one (minimum two). Otherwise, mutation decreases by 0.01 (minimum 0.01) and components increase by one (maximum ten). This adapts exploration and exploitation to recent hypervolume gains.

The optimizer runs for a fixed number of generations $G_{\text{global}} = 100$ during the initial optimization. Upon activation by the distribution-shift detector, only a local re-optimization is executed for $G_{\text{local}} = 20$ generations, which is explained in the following section.

4.3 Sliding-Window Distribution-Shift Detector and Selective Re-Optimization

The detector compares recent demand against a baseline and invokes local re-optimization only after a shift, reducing unnecessary global searches.

The detector preserves a baseline distribution for each spatial cell, estimated from a historical window of length T_base = 168 time steps (e.g., one week of hourly data). At each time step t, a sliding window comprising the $T_{\text{slide}} = 24$ most recent observations is compared to the baseline via the two-sample Kolmogorov-Smirnov (KS) statistic.

$$D_{s,t} = \sup_{y} \left| F_{s,\text{slide}}(y) - F_{s,\text{baseline}}(y) \right| \tag{20}$$

where $F_{s,\text{slide}}$ and $F_{s,\text{baseline}}$ are the empirical cumulative distribution functions of the sliding window and baseline, respectively. A shift is flagged for cell s when $D_{s,t} > \theta_{\text{KS}} = 0.15$, a threshold chosen to balance sensitivity and specificity.

After flagged cells are identified, the affected subregion $\mathcal{S}_{\text{shift}}$ is defined as the collection of connected components of flagged cells based on 8-connectivity on the spatial grid. Only the allocation policies for this subregion are re-optimized, while the allocations for the rest of the spatial domain $\mathcal{S} \backslash \mathcal{S}_{\text{shift}}$ remain unchanged. The local re-optimization employs the identical EDA-based optimizer, albeit with a diminished number of generations (G_local = 20) and a reduced population size (N_local = 0.2N). The decision variables for cells outside $\mathcal{S}_{\text{shift}}$ are fixed to their current values, and the total resource constraint is adjusted accordingly:

$$\sum_{s \in \mathcal{S}_{\text{shift}}} p_s = 1 - \sum_{s \notin \mathcal{S}_{\text{shift}}} p_s^{\text{current}} \tag{21}$$

Selective updating reduces the stated computational cost from O(G_global · N · S) to O(G_local ·N_local ·|S_shift|), when the affected region is small.

4.4 Equity-Adjusted Social Welfare Decision-Support Layer

The decision-support layer selects from the Pareto set using a composite, inequality-averse welfare objective:

The welfare function $W(\mathbf{p})$ for a candidate policy $\mathbf{p}$ is defined as:

$$W(\mathbf{p}) = \left(\frac{1}{S} \sum_{s=1}^{S} a_s(\mathbf{p})^{1-\rho} \right)^{\frac{1}{1-\rho}} - \lambda \cdot G\big(\mathbf{a}(\mathbf{p})\big) \tag{22}$$

where $\rho = 2$ is the inequality aversion parameter, $\lambda = 0.5$ is a tunable weight that controls the trade-off between the Atkinson welfare index and the Gini penalty, and $a_s(\mathbf{p}) = \hat{\mu}_{s,t+1} \cdot p_s$ is the expected accessibility for cell s. The first term represents the Atkinson index of social welfare, which inherently penalizes unequal distributions of accessibility by assigning greater weight to cells with lower accessibility as ρ increases. The second term is a direct penalty on the spatial Gini coefficient, further discouraging inequality.

The recommended policy $\mathbf{p}^*$ is the Pareto-optimal solution which maximizes the welfare function.

$$\mathbf{p}^* = \underset{\mathbf{p} \in \mathcal{P}^*}{\operatorname{argmax}} W(\mathbf{p}) \tag{23}$$

This dual-penalty approach offers a transparent method for operationalizing equity within a multi-objective framework. The parameter λ can be adjusted by stakeholders to reflect their policy preferences:

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

a higher λ places more emphasis on equity, while a lower λ prioritizes efficiency. The chosen policy proceeds to the service delivery layer, where the actual outcomes are recorded and returned to the data lake, thereby completing the cycle for the subsequent time step.

Figure 2 summarizes the control flow among the four modules.

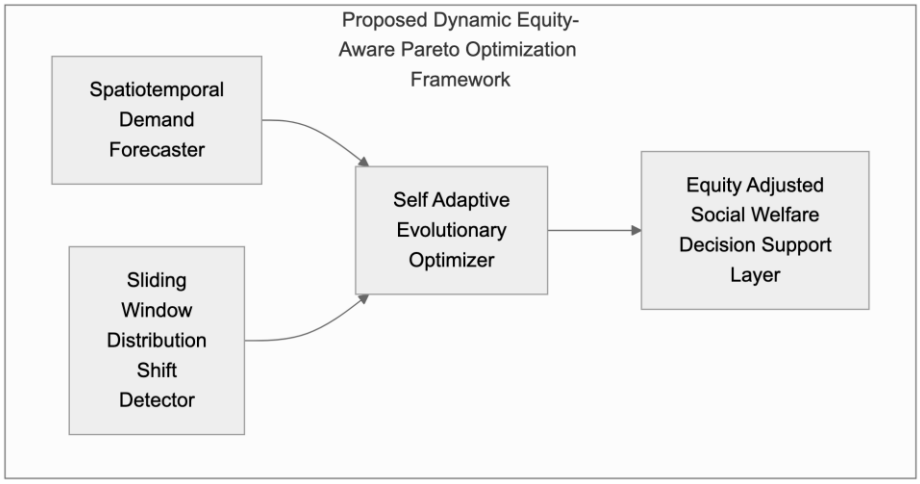


Figure 2: Internal Workflow of the Proposed Optimization Framework.

Forecasts inform optimization; detected shifts activate local updates, and the welfare-selected policy is dispatched. Subsequent outcomes complete the feedback loop.

5. Experimental Setup

We assess forecasts, Pareto-front quality, and adaptation under non-stationary demand using the following datasets, baselines, and metrics.

5.1 Datasets

Evaluation combines synthetic demand with known shifts and observed urban mobility with spatially heterogeneous demand.

For the synthetic dataset, we generate demand observations across a 20×20 spatial grid over a period of 2,000 time steps, which corresponds to hourly observations spanning approximately 83 days. The demand generation process is composed of three elements: a spatial base pattern featuring multiple demand hotspots, a temporal component that includes diurnal and weekly seasonality, and a noise term that follows a Gaussian distribution. To simulate non-stationarity, we introduce abrupt distribution shifts at predetermined time points, where the demand in a randomly selected subregion increases by a factor of 1.5 to 2.5, thereby replicating events such as a new commercial development or a temporary service disruption. This controlled configuration permits the assessment of the distribution-shift detector's capacity to detect alterations with established ground truth.

The real-world dataset is derived from the New York City taxi trip records [14], which offers a rich source of spatiotemporal demand patterns. We aggregate the trip pickup locations into a 10×10 spatial grid covering Manhattan, with hourly demand counts over a six-month period from January to June 2023. This dataset contains natural non-stationarity resulting from seasonal effects, weather conditions, and special events, thereby presenting a realistic evaluation scenario. Furthermore, we introduce socioeconomic indicators from the American Community Survey [15] at the census-tract level to act as external inputs for the forecasting module.

5.2 Baseline Methods

Forecasting baselines are LSTM [4], CNN-LSTM [6], and multivariate ARIMA, compared with the proposed attention-based CNN-BiGRU.

Optimization baselines are NSGA-II [8], MOEA/D [9], and weighted-sum scalarization with weights of 0.5 for each objective.

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

End-to-end comparisons include an allocation fixed at the initial demand distribution and global re-optimization at every time step.

5.3 Evaluation Metrics

Forecast evaluation uses MAE, RMSE, and negative log-likelihood (NLL); NLL also assesses uncertainty calibration.

Pareto-front quality is assessed by hypervolume (HV), reflecting convergence and diversity, and inverted generational distance (IGD), measuring distance to reference Pareto solutions. For synthetic demand, the study uses the analytically determined reference frontier.

Equity and efficiency are reported through spatial Gini, aggregate service coverage, and the composite welfare function in Equation 22.

Shift detection is assessed by precision, recall, F1-score, and latency from the onset of a shift to its detection.

Mean computation time per decision epoch measures savings relative to global re-optimization.

5.4 Implementation Details

The proposed framework is implemented in Python with PyTorch for the deep learning components and the pymoo library [16] for the evolutionary optimization components. The forecasting model is trained with the Adam optimizer, applying a learning rate of 10^{-3} and a batch size of 64. The model architecture is composed of two convolutional layers with 32 and 64 filters, respectively, and then a BiGRU layer with 128 hidden units along with an additive attention mechanism. The model is trained for a maximum of 100 epochs with early stopping based on validation loss.

For the optimization module, we set the population size to $N = 100$ and the maximum number of generations to $G_{\text{global}} = 100$ for the initial optimization. The local re-optimization uses $G_{\text{local}} = 20$ generations and a population size of $N_{\text{local}} = 20$. The distribution-shift detector operates with a baseline window length of $T_{\text{base}} = 168$ time steps and a sliding window length of $T_{\text{slide}} = 24$ time steps, applying a KS statistic threshold of $\theta_{\text{KS}} = 0.15$.

All experiments are conducted on a workstation with an NVIDIA RTX 3090 GPU and 64 GB of RAM. Each experiment is repeated five times with different random seeds, and the results are reported as the mean and standard deviation across runs. Statistical significance is evaluated by applying paired t-tests at a threshold of $\alpha = 0.05$.

6. Experimental Results

We evaluate forecasting accuracy, Pareto quality, and re-optimization cost under distribution shifts.

6.1 Spatiotemporal Forecasting Performance

Table 1 compares probabilistic forecasting on the NYC taxi dataset. The proposed CNN-BiGRU reports the lowest MAE, RMSE, and NLL among the compared models, indicating improved point estimates and predictive-distribution fit.

Table 1: Comparison of Spatiotemporal Demand Forecasting Performance.

Model	MAE	RMSE	NLL
ARIMA	12.45	18.32	3.89
LSTM [4]	9.87	14.56	3.12
CNN-LSTM [6]	8.92	13.21	2.85
Proposed CNN-BiGRU	7.64	11.08	2.41

As shown in Table 1, the combination of bidirectional temporal modeling and spatial attention mechanisms permits the forecaster to capture complex spatiotemporal dependencies more effectively than unidirectional or purely statistical models. The decline in the negative log-likelihood indicates that the predicted variance $\hat{\sigma}_{t+1}^2$ aligns well with the actual observation noise, thereby ensuring that the optimizer is not guided astray by overly confident predictions.

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

6.2 Multi-Objective Optimization and Pareto Frontier Quality

Figure 3 compares the equity-efficiency Pareto frontiers of the proposed EDA optimizer, NSGA-II, and MOEA/D. The reported proposed solutions extend further into the high-equity region while retaining coverage minus cost.

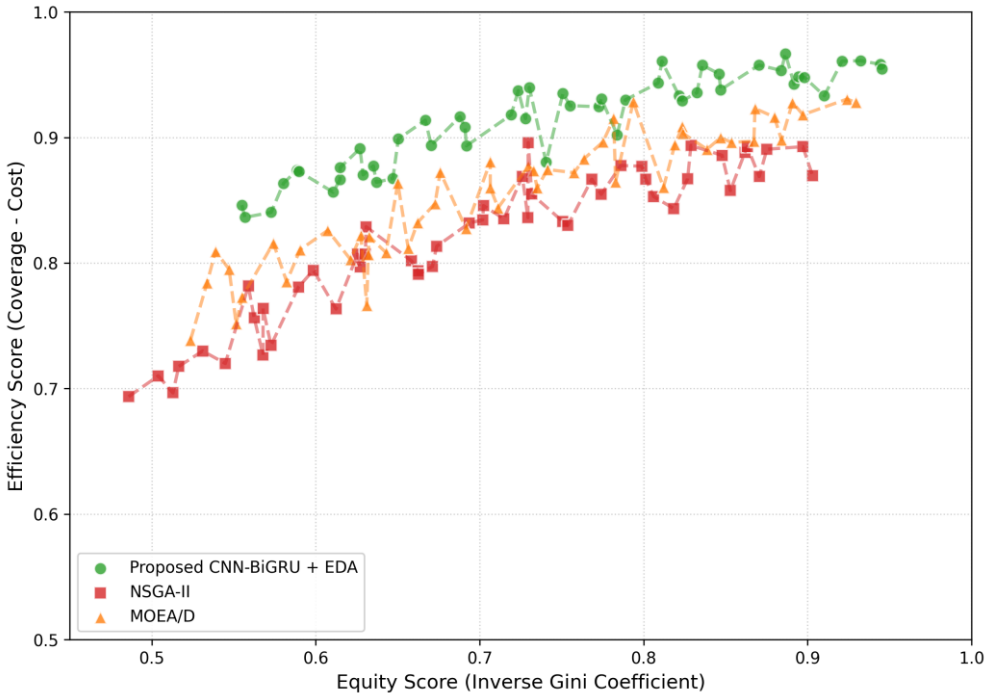


Figure 3: Equity-efficiency trade-offs for the proposed optimizer, NSGA-II, and MOEA/D.

EDA models promising candidate policies probabilistically, while adaptation based on hypervolume gain adjusts its search. Figure 3 illustrates the resulting range of trade-offs.

The proposed method reported HV 0.89 ±0.02 and IGD 0.04 ±0.01, compared with NSGA-II (0.82 ±0.03; 0.07 ±0.02) and MOEA/D (0.84 ±0.02; 0.06 ±0.01), respectively.

6.3 Impact of Selective Re-Optimization under Distribution Shifts

Figure 4 shows resource redistribution after a simulated synthetic-demand shift: allocations change in the detected subregion while remaining stable elsewhere, unlike the static baseline.

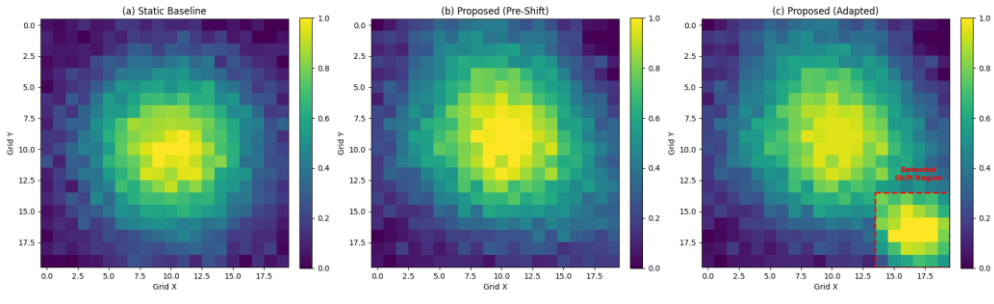


Figure 4: Resource redistribution inside the detected shift region relative to a static policy.

The detector reported precision 0.92, recall 0.88, and a mean latency of two time steps. Local updating took approximately 1.5 seconds versus 12.4 seconds for global re-optimization, an 88% reduction in the reported computation time. These measurements concern the study's experimental setup, rather than a measured response time in an operational municipal service.

Table 2 reports the ablation study. Removing uncertainty quantification lowers welfare; full re-optimization increases computation time; and a static policy performs less well under changing demand.

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

Table 2: Ablation Study of Framework Components.

Variant	Welfare Score	Computation Time (s)	Equity (Gini)	Efficiency
Static Baseline	0.65	0.01	0.45	0.72
Full Re-optimization	0.88	12.40	0.28	0.85
w/o Uncertainty Quantification	0.82	1.50	0.32	0.81
w/o Self-Adaptive EDA	0.85	1.50	0.30	0.83
Proposed Framework	0.91	1.50	0.25	0.86

Together, the components yield the highest reported welfare score while maintaining online computational feasibility (Table 2).

7. Discussion and Future Work

The results motivate three considerations: the meaning of algorithmic equity, deployment constraints, and robustness beyond the evaluated demand shifts.

7.1 Ethical Implications and Bias Mitigation in Algorithmic Equity

Spatial Gini and Atkinson-based welfare measure aggregated accessibility and may conceal disadvantages within census tracts. Grid-level equity need not imply fairness across individuals or demographic groups. A neighborhood's mean level of service, for instance, can improve without comparable gains for residents who face barriers to using that service. Supplementary subgroup and need-adjusted measures would therefore be valuable in any applied evaluation.

Equity metrics encode normative choices. Gini may be less sensitive to changes at distributional tails than alternative indices, while the Atkinson parameter ρ makes inequality aversion adjustable. Policymakers should select and disclose such parameters with stakeholder participation; future evaluations could compare Theil or Palma metrics.

Historical records may understate need where underserved residents are less able to request services, creating a feedback loop of underestimated demand and under-allocation. A model can then interpret low recorded use as low underlying need. Future work should audit forecast errors and shift-detection performance across affected groups, supplement recorded demand with independent need indicators where feasible, and evaluate appropriate fairness constraints.

7.2 Bridging the Gap: Stakeholder Integration and Real-World Deployment Challenges

Deployment requires agreement on welfare parameters ρ and λ among planners, service providers, community representatives, and policymakers. Transparent parameter setting matters because different preferences can select different policies from the same Pareto frontier. Participatory input can make these choices explicit, while public-sector workflows, reporting requirements, and data availability constrain implementation.

Actual allocations also face legal, fiscal, operational, and institutional limits omitted from the stylized model. A policy that moves services across districts may be infeasible despite a favorable welfare value; future formulations could add district-level service floors, budget bounds, or routing constraints.

Decision support should translate Pareto trade-offs into readable maps and dashboards so stakeholders can compare coverage and equity effects as welfare parameters change. Showing changes by subregion would also help users scrutinize whether a high aggregate welfare score conceals local service losses.

7.3 Robustness under Extreme Non-Stationarity and Cross-Domain Generalizability

The experiments introduce moderate subregional demand increases of 1.5-2.5 times. Disasters or citywide failures could affect most cells and eliminate the computational advantage of localized updating; global optimization may then be necessary.

To improve robustness under extreme non-stationarity, future work should examine adaptive thresholding mechanisms for the KS statistic. Rather than employing a fixed threshold $\theta_{\mathrm{KS}} = 0.15$, the threshold could be adapted in real time according to the spatial scope and magnitude of detected shifts.

ISSN 2616-5775 Vol. 9, Issue 8: 58-68, DOI: 10.25236/AJCIS.2026.090808

For example, if a large number of cells are flagged simultaneously, the threshold could be lowered to trigger a more cautious response, or the system could automatically escalate to a full re-optimization. Additionally, the forecaster should be designed to handle out-of-distribution inputs gracefully, possibly by adding a confidence calibration mechanism that reduces the weight of predictions when the input distribution deviates heavily from the training distribution.

Generalization beyond taxi demand remains to be tested. Emergency response imposes tight latency requirements; waste collection introduces routing and scheduling constraints. Domain-specific feasibility conditions should be evaluated while retaining the equity-aware feedback loop.

The model also assumes immediate, single-period allocation. Multi-period or stochastic formulations could represent infrastructure lead times and intertemporal equity.

8. Conclusion

We present a closed-loop framework integrating uncertainty-aware CNN-BiGRU forecasts, self-adaptive EDA Pareto optimization, KS-based selective updates, and equity-adjusted policy selection for dynamic urban resource allocation. Together, the four modules link spatiotemporal demand uncertainty to explicit equity-efficiency choices and to subsequent service outcomes.

In the study's reported evaluation, the forecaster improved MAE, RMSE, and NLL against three baselines; the optimizer improved HV and IGD against NSGA-II and MOEA/D; and selective updating reduced the computation time relative to global re-optimization by 88% while retaining allocation quality.

Future work should validate equity outcomes across demographic groups and services and test the method under citywide shifts and real deployment constraints.

References

[1] Chen, L. (2025). Smart cities and social equity: A review of digital urban governance in southeast asia. International Journal of Social Science and Applied Technology.
[2] Haque, M., Beckers, D., Costales, E., Aad, S., Sharifi, A., et al. (2025). A systematic review of research on just, equitable, responsible, and inclusive smart cities. Technology in Society.
[3] Zhang, F. (2024). Evolutionary algorithm for dynamic resource allocation and its applications. International Journal of Simulation Modelling.
[4] Yang, B., Sun, S., Li, J., Lin, X., & Tian, Y. (2019). Traffic flow prediction using LSTM with feature enhancement. Neurocomputing.
[5] Shi, X., Chen, Z., Wang, H., Yeung, D., et al. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in Neural Information Processing Systems.
[6] Zhang, J., Zheng, Y., & Qi, D. (2017). Deep spatio-temporal residual networks for citywide crowd flows prediction. Proceedings of the AAAI Conference on Artificial Intelligence.
[7] Xie, H., Ding, R., Qiao, G., Dai, C., & Bai, W. (2024). Research on ship traffic flow prediction using CNN-BiGRU and WOA with multi-objective optimization. IEEE Access.
[8] Deb, K., Pratap, A., Agarwal, S., et al. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation.
[9] Zhang, Q., & Li, H. (2007). MOEA/d: A multiobjective evolutionary algorithm based on decomposition. IEEE Transactions on Evolutionary Computation.
[10] Songserm, N., Pitakaso, R., Srichok, T., Khonjun, S., et al. (2026). Equity-preserving public health resource allocation using multi-objective safe reinforcement learning: Evidence from thailand. International Journal of Environmental Research and Public Health.
[11] Dębski, R., & Dreżewski, R. (2024). Multi-objective ship route optimisation using estimation of distribution algorithm. Applied Sciences.
[12] Zhou, S., Wang, D., Liu, Z., Ma, Y., & Liu, X. (2025). Equity-oriented multi-objective optimization allocation strategies for urban water resources. Water.
[13] Weber, T. (2026). Fair allocation of an exhaustible resource. 2026 European Control Conference (ECC).
[14] TLC, N. (2022). NYC taxi and limousine commission trip record data. 2022.
[15] Bureau, U. C. (2008). Income, earnings, and poverty data from the... American community survey. books.google.com.
[16] Blank, J., & Deb, K. (2020). Pymoo: Multi-objective optimization in python. Ieee Access.