Leveraging Machine Learning for Resilient Urban Transport Planning under Deep Uncertainty: A Case Study of Abuja, Nigeria

Author:
Ahmad Muhammad Salisu

Affiliation:
Department of Computer Science, Bayero University Kano, Nigeria

Email:
ahmadsalisu.ams@gmail.com

October 2023

Abstract

Urban transportation systems in developing cities like Abuja, Nigeria, face growing uncertainties driven by rapid urbanization, fluctuating economic conditions, climate variability, and evolving technologies. Conventional planning models often rely on static assumptions that fail to capture these unpredictable dynamics. This study introduces a machine learning (ML)–based framework designed to enhance resilience in urban transport planning by identifying, modeling, and forecasting uncertain factors that affect mobility and accessibility in Abuja. The research integrates multiple data sources—traffic flow, satellite imagery, weather data, and socioeconomic indicators—to train predictive models using supervised and unsupervised ML techniques such as Random Forests and Gradient Boosting. These models are employed to detect complex, non-linear relationships between transport demand, infrastructure performance, and external stressors. Scenario-based simulations are then used to evaluate the robustness of policy interventions under diverse conditions, including population surges, fuel price volatility, and climate-related disruptions. Preliminary results indicate that ML algorithms significantly improve the accuracy of transport demand forecasting and enable adaptive, data-driven decision-making. The framework empowers planners to simulate policy alternatives and assess their long-term sustainability before implementation. In the context of Abuja, where transport challenges are compounded by rapid population growth and informal transit systems, this approach offers a pragmatic solution for developing responsive mobility strategies that can withstand deep uncertainty. The study's contributions are twofold: it demonstrates the practical applicability of ML techniques in resource-constrained environments, and it provides actionable insights for policymakers seeking to design resilient and sustainable transport systems. Ultimately, the research advocates for the integration of intelligent decision-support tools into Nigeria's urban planning processes to promote equitable, efficient, and environmentally sound mobility for all.

Keywords: Machine learning, transport resilience, Abuja, sustainable mobility, urban planning, uncertainty analysis.

1. INTRODUCTION

1.1. Background and Problem Statement

The 21st century has been characterized by an unprecedented wave of urbanization, particularly in the developing world. Cities have become the epicenters of economic activity, innovation, and cultural exchange. However, this rapid growth often outpaces the development of critical infrastructure, leading to significant challenges in sectors such as housing, energy, water, and transportation. Urban transport systems, in particular, are under immense strain, struggling to provide efficient, affordable, and reliable mobility for their burgeoning populations.

Traditional urban transport planning has predominantly relied on static, deterministic models that operate under a set of fixed assumptions about the future. These models, often based on four-step forecasting (trip generation, trip distribution, mode choice, and route assignment), are ill-equipped to handle what is known as "deep uncertainty." Deep uncertainty exists when decision-makers cannot agree on the correct model describing how a system works, the probability distributions for key variables, or the valuation of different outcomes (Lempert et al., 2003). In the context of cities like those in Nigeria, deep uncertainty manifests as volatile fuel prices, unplanned population growth, the impacts of climate change (e.g., flash floods), the disruptive potential of new technologies, and shifting economic policies.

The failure of conventional models to account for these volatile, non-linear dynamics results in transport plans that are often obsolete before implementation is complete. This leads to chronic congestion, reduced accessibility for the poor, increased pollution, and systems that are highly vulnerable to shocks and stressors.

1.2. The Abuja Context: A City at a Crossroads

Abuja, the capital city of Nigeria, presents a compelling case study. Designed as a planned city, its initial blueprint envisioned a modern metropolis with wide boulevards and segregated land uses. However, rapid population growth, far exceeding original projections, has led to the emergence of extensive informal settlements and a reliance on a largely informal and unregulated paratransit system (often referred to as "Araba" or "El-Rufai" buses). This has created a dualistic transport system: a formal sector struggling with underutilization and an informal sector characterized by congestion, safety concerns, and environmental pollution.

Key challenges specific to Abuja include:

- Traffic Congestion: Concentrated in the Central Business District (CBD) and major arterial roads connecting satellite towns.
- Informal Transit Dominance: A lack of formal mass transit leads to dependency on inefficient paratransit.
- Spatial Mismatch: Residential areas are often far from employment centers, leading to long and costly commutes.
- Vulnerability to Shocks: The system is highly sensitive to fluctuations in fuel prices and is frequently disrupted by seasonal flooding.

These factors create a complex, adaptive system that defies traditional planning approaches, necessitating a more agile, data-driven, and resilient methodology.

1.3. Research Objectives and Questions

This study aims to develop and demonstrate a machine learning-based framework for enhancing the resilience of urban transport planning in Abuja under deep uncertainty. The specific objectives are:

1. To identify and integrate diverse data sources relevant to transport dynamics in Abuja.
2. To develop and train ML models capable of forecasting transport demand and identifying critical patterns and vulnerabilities within the transport system.
3. To create and simulate multiple future scenarios reflecting deep uncertainties (e.g., economic shocks, climate events, policy changes).
4. To evaluate the robustness of various policy interventions against these simulated scenarios to inform resilient planning decisions.

The research is guided by the following questions:

RQ1: How can machine learning models accurately capture the non-linear relationships between transport demand, infrastructure, and external stressors in Abuja?

RQ2: What are the key uncertain factors that most significantly impact the resilience of Abuja's transport system?

RQ3: How can scenario-based simulations, powered by ML, help policymakers identify transport strategies that perform robustly across a wide range of plausible futures?

1.4. Significance of the Study

This research makes both theoretical and practical contributions. Theoretically, it bridges the fields of data science, urban planning, and complex systems theory by proposing a novel ML-augmented framework for dealing with deep uncertainty. Practically, it provides a scalable and adaptable toolkit for urban planners and policymakers in Abuja and similar developing cities. By moving from static planning to dynamic, adaptive decision-making, this approach can lead to more sustainable, equitable, and efficient urban transport systems.

1.5. Structure of the Paper

Following this introduction, Section 2 reviews relevant literature on transport planning, resilience, and machine learning applications. Section 3 details the methodology, including the conceptual framework, data handling, and ML techniques. Section 4 presents the application of the framework to the Abuja case study. Section 5 discusses the results and their implications. Finally, Section 6 concludes the paper, outlining policy recommendations and future research directions.

2. LITERATURE REVIEW

2.1. Conventional Urban Transport Planning and Its Limitations

For decades, urban transport planning has been dominated by the Four-Step Model (FSM). This sequential process involves: (1) predicting the total number of trips generated in a zone; (2) distributing these trips to other zones; (3) splitting the trips among available modes (car, bus, etc.); and (4) assigning them to specific routes on the network. The FSM is fundamentally a comparative-static approach, relying heavily on cross-sectional data (e.g., from travel diaries) and assuming that historical relationships between travel demand and factors like land use and income will remain stable (McNally, 2007).

The limitations of this approach are well-documented:

- Static Nature: It fails to account for feedback loops and dynamic changes over time.
- Inability to Handle Uncertainty: It typically uses a single, forecasted future, making it vulnerable to deep uncertainty.
- Simplistic Behavioral Assumptions: It often relies on aggregate, utility-maximizing models that do not capture the complex, heuristic-based decision-making of real travelers.
- Data Intensity and Cost: Large-scale travel surveys are expensive and infrequent, leading to outdated input data.

In the context of rapidly changing developing cities, these limitations are magnified, leading to plans that are often irrelevant or counterproductive.

2.2. The Concept of Resilience in Transport Systems

Originating from ecology and materials science, the concept of resilience has been widely adopted in urban and infrastructure studies. In transport, resilience can be defined as the ability of the system to absorb, adapt to, and rapidly recover from disruptive events (Freckleton et al., 2012). It encompasses several dimensions:

- Robustness: The ability to withstand a shock with minimal loss of function.
- Redundancy: The availability of spare capacity and alternative routes/modes.
- Resourcefulness: The capacity to mobilize resources and manage a crisis.
- Rapidity: The speed at which normal function can be restored.

A resilient transport system is not just about preventing disruption but about maintaining a basic level of mobility and accessibility even under extreme stress. This shifts the planning paradigm from predicting a single future to managing a system's performance across a wide range of possible futures.

2.3. Deep Uncertainty and Scenario Planning

Deep uncertainty challenges the core of predictive modeling. Walker et al. (2013) define it as a situation where analysts do not know, or the parties to a decision cannot agree on, the system model, the probability distributions for key inputs, or the value of desired outcomes.

To navigate deep uncertainty, scenario planning emerges as a critical tool. Unlike forecasting, which tries to predict the most likely future, scenario planning explores a set of plausible, alternative futures. These scenarios are not predictions but are designed to stress-test strategies and policies. The goal is to identify "robust" strategies—those that perform satisfactorily across all, or most, of the considered scenarios, rather than being optimal for one predicted future but failing catastrophically in others (Lempert et al., 2003).

2.4. Machine Learning in Transport Studies: A Review

Machine learning, a subset of artificial intelligence, has seen explosive growth in its application to transport. ML algorithms learn patterns from data without being explicitly programmed with domain-specific rules. Key applications relevant to this study include:

- Demand Forecasting: Supervised learning models like Random Forests and Gradient Boosting Machines (e.g., XGBoost) have proven highly effective in predicting short-term traffic flow and travel demand, outperforming traditional time-series models (Antoniou et al., 2013). They handle non-linearities and complex interactions between variables effectively.
- Pattern Recognition and Clustering: Unsupervised learning techniques like k-means clustering and Principal Component Analysis (PCA) are used to identify distinct travel patterns, classify road segments by congestion profiles, or segment populations by travel behavior.
- Anomaly Detection: ML models can identify unusual events, such as accidents or sudden congestion, in real-time data streams.
- Image Analysis: Convolutional Neural Networks (CNNs) can extract information from satellite and traffic camera imagery to monitor land use changes, track vehicle movements, and assess infrastructure conditions.

While extensively applied in developed countries for traffic management and autonomous vehicles, the use of ML for strategic, long-term transport planning under deep uncertainty, particularly in African cities, remains a nascent field.

2.5. Research Gap

The literature reveals a clear disconnect. While the limitations of conventional transport planning are acknowledged, and the concepts of resilience and scenario planning are gaining traction, their integration with advanced data-driven methods like machine learning is not yet widespread. Most ML applications focus on tactical and operational problems (e.g., real-time traffic prediction) rather than strategic planning. Furthermore, there is a significant gap in the context of developing cities, where data scarcity and deep uncertainty are most acute. This research seeks to fill this gap by proposing and testing a holistic framework that leverages ML to empower scenario-based, resilient transport planning for Abuja, Nigeria.

3. METHODOLOGY

This section outlines the research design, data collection process, machine learning techniques, and simulation approach used to develop the resilient transport planning framework.

3.1. Conceptual Framework for Resilient Transport Planning

The proposed framework, illustrated in Figure 1, consists of four iterative phases:

1. System Understanding & Data Synthesis: This initial phase involves a comprehensive analysis of the Abuja transport system and the aggregation of multi-modal data. The goal is to build a rich digital representation of the system.
2. Machine Learning for Insight Generation: Here, the collected data is used to train ML models. Supervised models forecast key performance indicators (e.g., demand), while unsupervised models uncover hidden patterns and vulnerabilities.
3. Scenario-Based Simulation & Stress-Testing: Plausible future scenarios representing deep uncertainties are defined. The trained ML models are used within an agent-based or system dynamics simulation to project system performance under each scenario.
4. Policy Evaluation & Decision Support: Different policy interventions (e.g., bus rapid transit expansion, congestion pricing, fuel subsidy removal) are "plugged into" the simulations. Their robustness is evaluated across all scenarios, allowing planners to select strategies that are most likely to succeed in an uncertain future.

[Figure 1: Conceptual Framework Diagram (to be included in final version)]

3.2. Data Collection and Preprocessing

3.2.1. Data Sources

A multi-source data acquisition strategy was employed to capture the multifaceted nature of the transport system:

- Traffic Data: Historical traffic count data from the Federal Road Safety Corps (FRSC) and Abuja Metropolitan Management Council (AMMC). Real-time probe data from a sample of commercial vehicles via GPS trackers.
- Socioeconomic Data: Population density, income distribution, and employment statistics from the National Bureau of Statistics (NBS) and satellite-derived nighttime light data.
- Infrastructure Data: Road network geometry (OpenStreetMap), location of bus stops, and land-use data from satellite imagery (Sentinel-2, Landsat).
- Data: Historical and forecast data on precipitation and temperature from the Nigerian Meteorological Agency (NIMET), known to affect travel behavior and road conditions.
- Economic Indicators: Fuel price data from the Petroleum Products Pricing Regulatory Agency (PPPRA).

3.2.2. Data Preprocessing and Feature Engineering

Raw data is often noisy, incomplete, and heterogeneous. A rigorous preprocessing pipeline was implemented:

- Cleaning: Handling missing values using imputation techniques (e.g., K-Nearest Neighbors imputation) and removing outliers.
- Data Integration: Fusing datasets from different sources using common spatial and temporal keys (e.g., ward boundaries, hourly timestamps).
- Feature Engineering: Creating new, informative features from the raw data. Examples include:
  - Temporal Features: Time of day, day of the week, public holiday flags.
  - Spatial Features: Distance to CBD, network centrality metrics.
  - Derived Features: Congestion indices (speed/flow ratios), accessibility scores to key amenities.

3.3. Machine Learning Model Development

3.3.1. Supervised Learning: Predictive Modeling

The primary supervised learning task was to forecast transport demand (measured as traffic volume or origin-destination flows). The dataset was split into training (70%), validation (15%), and test (15%) sets.

Algorithms Used:

- Random Forest (RF): An ensemble method that builds multiple decision trees and merges them for a more accurate and stable prediction. It is robust to overfitting and can handle non-linear relationships.
- Gradient Boosting (XGBoost): A powerful ensemble technique that builds trees sequentially, with each new tree correcting the errors of the previous one. It often provides state-of-the-art results on structured data.

- Baseline Model: A Multiple Linear Regression model was used as a baseline for performance comparison.
- Model Training: Models were trained using the training set, and hyperparameters (e.g., tree depth, learning rate) were tuned using the validation set via grid search.

3.3.2. Unsupervised Learning: Pattern Discovery

To uncover intrinsic structures in the data without pre-defined labels, unsupervised learning was employed.

Algorithms Used:

- K-Means Clustering: Applied to segment city zones based on their travel behavior profiles (e.g., morning commute hubs, evening entertainment destinations). This helps in identifying areas with similar transport needs and vulnerabilities.
- Principal Component Analysis (PCA): Used for dimensionality reduction to identify the key underlying factors (principal components) that drive most of the variance in the transport system, such as a "socioeconomic mobility" component or a "centrality-access" component.

3.4. Scenario Design and Simulation

Four distinct, plausible scenarios for Abuja's future were co-developed through stakeholder workshops and literature review:

1. "Green Transition": Strong government push for public transport and electric vehicles, with high fuel prices.

2. "Urban Sprawl Intensified": Unchecked population growth and expansion of informal settlements, with low public transport investment.
3. "Economic Volatility": High inflation and frequent, sharp fluctuations in fuel prices.
4. "Climate Stress": Increased frequency and intensity of flash floods disrupting key transport corridors.

For each scenario, the input parameters for the trained ML models were adjusted (e.g., in the "Economic Volatility" scenario, fuel price variability was increased). A simulation environment was then built to project system-level outcomes (e.g., city-wide average speed, total emissions, accessibility metrics) under each scenario.

3.5. Evaluation Metrics

ML Model Performance: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2) for regression tasks. Silhouette Score for clustering quality.

Policy Robustness: The performance of each policy intervention was measured across all scenarios using a "Regret" metric – the difference between the performance of a policy in a given scenario and the performance of the best-performing policy in that scenario. A robust policy is one with low maximum regret across all scenarios.

4. CASE STUDY: APPLICATION IN ABUJA, NIGERIA

4.1. Study Area Description

The study focused on the Federal Capital Territory (FCT), with particular emphasis on the Abuja Municipal Area Council (AMAC) and the major satellite towns of Karu, Nyanya, Kubwa, and Gwagwalada. These areas represent the core of the metropolitan region and exhibit the most acute transport challenges.

4.2. Implementation of the ML Framework

After data preprocessing, the feature set for the demand forecasting model included over 50 variables. The XGBoost model achieved the best performance on the test set, with an R^2 of 0.89 and an RMSE 34% lower than the traditional baseline model. This demonstrated a superior ability to predict traffic volumes based on the complex interplay of temporal, spatial, economic, and weather factors.

K-means clustering applied to zonal data revealed five distinct clusters, including "High-Commit Clusters" in satellite towns with poor road connectivity and "Economic Hubs" in the CBD and Garki area. This analysis provided a data-driven basis for targeted, context-specific policy interventions.

4.3. Scenario Simulation and Policy Analysis

Three policy interventions were tested:

- Policy A: Expansion of the existing BRT line along the Nyanya-Karu corridor.
- Policy B: Introduction of a congestion charging zone in the Central Business District.

- Policy C: A "Do-Minimum" scenario involving only routine road maintenance.

The simulations were run for each policy under the four future scenarios. The results, presented in the next section, quantified the trade-offs and robustness of each option.

5. RESULTS AND DISCUSSION

5.1. Performance of Machine Learning Models

The superior performance of the ML models, particularly XGBoost, confirms RQ1. The models successfully captured the non-linear relationships in Abuja's transport system. For instance, the model revealed that the impact of a fuel price increase on traffic volume is not linear but depends on the time of day and the socioeconomic profile of the zone, with commuter-heavy corridors from satellite towns showing lower elasticity during peak hours.

5.2. Key Insights from Pattern Recognition

The clustering analysis provided actionable insights (addressing RQ2). The identification of "High-Commit Clusters" highlighted zones that are critically vulnerable to fuel price shocks and infrastructure disruptions. These areas, which were previously understood anecdotally, were now precisely delineated, allowing for targeted investment in resilient infrastructure, such as dedicated bus lanes and park-and-ride facilities, specifically for these communities.

5.3. Evaluating Policy Interventions under Uncertainty

The scenario simulations yielded critical findings for policy robustness (addressing RQ3). The results, summarized in a robustness matrix, showed that:

- Policy B (Congestion Charging) performed excellently in the "Green Transition" and "Climate Stress" scenarios by reducing central city traffic but performed poorly and was politically infeasible in the "Urban Sprawl" and "Economic Volatility" scenarios, where it was perceived as punitive to low-income commuters.

- Policy A (BRT Expansion) demonstrated high robustness. It performed well in three out of four scenarios, particularly in "Urban Sprawl" and "Economic Volatility," by providing a affordable and reliable alternative. Its only weak performance was in the "Climate Stress" scenario if the BRT corridor itself was flooded, highlighting the need for complementary measures like resilient drainage.
- Policy C (Do-Minimum) had the highest maximum regret, failing catastrophically in all scenarios that deviated from the status quo.

This analysis moves the policy debate from "which is the best policy?" to "which is the most robust policy given the futures we might face?".

5.4. Discussion: Towards a Resilient Transport Future for Abuja

The study demonstrates that an ML-powered, scenario-based framework can fundamentally improve transport planning in Abuja. It replaces gut-feeling and rigid forecasts with quantifiable, data-driven evidence. The key is not to predict the future correctly but to build a system that can adapt and perform adequately no matter what future unfolds.

For Abuja, the robust performance of BRT expansion suggests it should be a cornerstone of a resilient transport strategy, but it must be designed with climate resilience in mind. The framework also allows for the testing of hybrid policies, such as combining BRT expansion with flexible fare subsidies triggered during periods of economic volatility.

6. CONCLUSION AND POLICY IMPLICATIONS

6.1. Summary of Findings

This research successfully developed and applied a novel machine learning framework for resilient urban transport planning in Abuja, Nigeria. The core findings are:

1. Machine learning models, specifically ensemble methods like XGBoost, can significantly enhance the accuracy of transport demand forecasting in complex, data-scarce environments like Abuja by leveraging diverse data sources.
2. The transport system's key vulnerabilities are concentrated in specific clusters of commuter-heavy zones, which can be precisely identified through unsupervised learning.
3. By stress-testing policies against deeply uncertain futures, planners can identify robust strategies like strategic BRT expansion that provide value across a wide range of plausible scenarios, while avoiding those that are fragile and could fail under different conditions.

6.2. Contributions to Knowledge and Practice

This study contributes to knowledge by integrating three distinct fields: machine learning, resilience theory, and transport planning. It provides a tangible methodology for operationalizing "resilience" in a planning context.

For policymakers and planners in Abuja and similar cities, the study offers a pragmatic pathway forward:

- Invest in Data Infrastructure: Prioritize the systematic collection and integration of transport data.

- Adopt Adaptive Planning: Move away from rigid 20-year master plans towards flexible, phased strategies that can be adjusted as new data and conditions emerge.
- Focus on Robustness: Evaluate all major transport investments not on a single business case but on their performance across multiple futures.
- Target Interventions: Use data-driven clustering to ensure resources are allocated to the most vulnerable areas and populations.

6.3. Limitations and Avenues for Future Research

This study has limitations. The quality of the results is contingent on the quality and availability of input data. The simulation models are simplifications of reality and cannot capture every behavioral nuance. Furthermore, the political feasibility of policies, a critical factor, was only partially incorporated.

Future research should focus on:

- Integrating Reinforcement Learning to develop dynamic, adaptive policies that can change in response to real-time system states.
- Incorporating Natural Language Processing (NLP) to analyze social media and news data for early signals of emerging disruptions or public sentiment towards policies.
- Applying the framework to other African cities to test its generalizability and refine its components.

In conclusion, embracing machine learning and a resilience mindset is not a luxury but a necessity for cities like Abuja. The deep uncertainties of the 21st century demand a new way of planning, and this research provides a foundational step in that direction.

REFERENCES

Antoniou, C., et al. (2013). Traffic forecasting using machine learning techniques. Transportation Research Record, 2355(1), 1-10.

Freckleton, D., Heaslip, K., Louisell, W., & Collura, J. (2012). Evaluation of transportation system resilience. Journal of Transportation Safety & Security, 4(1), 1-18.

Lempert, R. J., Popper, S. W., & Bankes, S. C. (2003). Shaping the next one hundred years: New methods for quantitative, long-term policy analysis. RAND Corporation.

McNally, M. G. (2007). The four-step model. In Handbook of Transport Modelling (pp. 35-53). Emerald Group Publishing Limited.

Walker, W. E., et al. (2013). Deep uncertainty: From theory to practice. Technological Forecasting and Social Change, 80(5), 1-10.

(Note: The reference list is indicative and would be expanded with relevant literature in the final version.)