EAI Endorsed Transactions
on Internet of Things
Research Article EAI.EU

Quantum-Classical Adaptive Navigation: QAOA-PPO Framework for GNSS-Denied Urban Environments

Gabriel Silva-Atencio[1,*]

[1]Universidad Latinoamericana de Ciencia y Tecnología (ULACIT), San José, Costa Rica

Abstract

This study shows a mixed quantum-classical guidance system that combines an eight-layer Quantum Approximate Optimization Algorithm (QAOA) with Proximal Policy Optimization (PPO) reinforcement learning for cities that do not have GNSS. The design changes the settings of the quantum circuit on the fly based on real-time data from the environment. This makes noise in cities less of a static mistake source and more of an optimization constraint. After testing on 47.3 kilometers of GNSS-denied paths in three cities using a three-tier DARPA QuANET-aligned protocol and NIST-traceable equipment, the results show that the average positional accuracy is 0.15 ± 0.03 meters, which is 70% better than high-grade GPS/INS systems. With adaptive correction, the framework keeps 98.2% of its coherence under ISO 16750-3 shaking and 95.3% of its operational reliability during 24-hour signal rejection. It also reduces drift by 82% compared to LiDAR-SLAM over a kilometer. It is statistically significant ($F(3,1196) = 87.3$, $p < 0.001$, $\cdot^2 = 0.38$), and the effect sizes are big. The system passes SAE Level 4 power limits (45.2W), goes beyond DARPA QuANET 2025 goals, and gets Hybrid Readiness Level 7 approval. Each unit is expected to cost $210,000, and it creates the first policy-ready quantum navigation platform for smart city application.

Keywords: Autonomous vehicles, GNSS-denied navigation, NIST validation, quantum approximate optimization algorithm, reinforcement learning, urban positioning

Received on 24 November 2025, accepted on 02 April 2026, published on 16 April 2026

doi: 10.4108/eetiot.11077

*Corresponding author. Email: gsilvaa468@ulacit.ed.cr

1. Introduction

The increasing number of high-rise buildings, underground transit networks, and dense Internet of Things (IoT) deployments seriously threatens the integrity of Global Navigation Satellite Systems (GNSS). Modern societies are currently facing an unmatched crisis of reliability in urban areas. These structural and electromagnetic problems cause serious multipath transmission, signal loss, and chronic non-line-of-sight situations. For this reason, standard GNSS systems are becoming less useful for applications that require safety [1, 2]. These types of positioning errors, which can exceed thirty meters in deep urban canyons, differ considerably from the sub-decimeter accuracy required by SAE Level 4 certified autonomous vehicles, emergency response logistics, and precise infrastructure monitoring [3, 4].

This decline is not just a minor change; it is a fundamental collapse of the physical ideas on which satellite positioning is based. In cities, GNSS signals from reliable data sources are turned into random noise. In metropolitan cores, direct line-of-sight availability drops below 40%, and errors caused by multipaths do not follow a normal Gaussian distribution, which means that standard Kalman filtering methods do not work.

Multisensor fusion designs that combine Light Detection and Ranging (LiDAR), visual odometry, and inertial measurement units (IMUs) have been used in traditional avoidance methods. However, these traditional methods have two major problems that are becoming more and more untenable: they can't handle the huge amounts of data that are created in crowded cities; and they make huge amounts of mistakes when GNSS signals

go out, with drift rates of up to one meter per kilometer in tunnels and underground facilities [5, 6]. As the world gets more complicated, the amount of work that needs to be done on the computer increases very quickly because feature extraction methods must deal with point clouds that are getting denser while still having to work in real time with delay levels below 50 milliseconds.

Even though ORB-SLAM3 and similar frameworks have improved algorithms, simultaneous localization and mapping (SLAM) implementations are still prone to perceptual aliasing in cities with lots of buildings that look the same and to failing completely in visually poor conditions like tunnels, heavy rain, or at night [3, 5]. These issues are not just minor technical matters, but fundamental design decisions. Traditional navigation systems still rely on outside signals or long-lasting features in the environment, neither of which work reliably in the operational conditions that define modern smart city infrastructure. It is hard to make truly independent urban environments because of the slow processing speed and easy signal loss. This is because operational reliability standards require failure rates below ten to minus nine per hour of operation.

In recent years, quantum inertial navigation systems (Q-INS) have become a potentially transformative concept. They use quantum superposition and entanglement to achieve signal-independent tracking with minimal error. Technologies like cold-atom interferometers (CAI) and nitrogen-vacancy (NV) center magnetometers have shown that they can measure with a centimeter of accuracy over kilometer-scale paths, which means they have the potential to be more stable and accurate than ever before [7-9]. This is possible with CAI because they use the wave-particle duality of ultracold atomic groups to create superpositions that are sensitive to accelerations and rotations but not at all sensitive to electromagnetic interference from the outside.

Instead, NV centers use flaws in diamond lattice structures that are like atoms and can be set up and read out visually. This makes the magnetic field sensitive at the nanotesla level, which is enough for lane-level classification against magnetic anomaly maps that have already been studied [10, 11]. Gravimetric sensitivity below ten to the minus nine g per root hertz has been demonstrated in the lab, along with magnetic field resolution sufficient to detect individual electron spins with NV centers. These results lay the theoretical groundwork for decimeter-level absolute positioning without external references.

Despite this, a thorough study of the existing literature shows a huge gap in translation. As shown in Table 1 and shown visually in Figure 1, less than 12% of quantum navigation studies go beyond controlled lab settings to look at what needs to be done for practical urban usage [12, 13]. The difference between what is possible in the lab and what is possible in cities is caused by three problems that are still not fixed. These problems can be put together and called the "quantum urban navigation gap."

Table 1. Quantum Navigation Research Translational Gap: Laboratory Promise vs. Urban Reality

Metric	Laboratory Performance (Typical)	Urban Deployment Requirement	Critical Challenge
Accuracy (RMSE)	0.01 – 0.1 m	≤ 0.25 m (SAE Level 4)	Vibration/EMI Decoherence
Coherence Time (T_2)	> 1 s	> 100 ms (operational)	Environmental Noise
Power Consumption	~100s of W (benchtop)	≤ 50 W (vehicular)	Cryogenic Integration
Environmental Validation	Controlled optical tables	Real-world urban canyons	< Twelve percent of studies include urban testing

The first challenge is the lack of order in the surroundings. Heavy electromagnetic interference (five to one hundred megahertz), ground-borne vibrations from zero point one to one hundred hertz, and big temperature differences from -20 to +50 degrees Celsius are all types of noise that are unique and bad for electronics. These stresses cause quantum decoherence to happen quickly, which makes sensors work much worse than they did in the lab [10, 14]. For example, electromagnetic radiation at megahertz frequencies directly couples to qubit energy levels, causing Rabi oscillations that destroy superposition states; mechanical vibrations at frequencies below one hundred hertz cause phase noise in atom interferometers through Doppler shifts and strain-induced splitting in solid-state systems; and changes in temperature change the spin Hamiltonian parameters and change the rates of relaxation through temperature-dependent phonon coupling. These physical processes are well understood but hard to make technologically.

As you move, the coherence time parameter T_2 can go from seconds on vibration-isolated optical tables to milliseconds in moving vehicles. This is a three-orders-of-magnitude drop in the amount of time that quantum superpositions can last, which makes it hard for quantum sensors to be used outside of libraries. As well as the engineering challenges that come with mobile deployment that are not often talked about in basic quantum sensing literature, these include keeping the ultra-high vacuum conditions needed for cold-atom operation and keeping the temperature stable within plus or minus 0.1 Kelvin for NV center spin-state initialization.

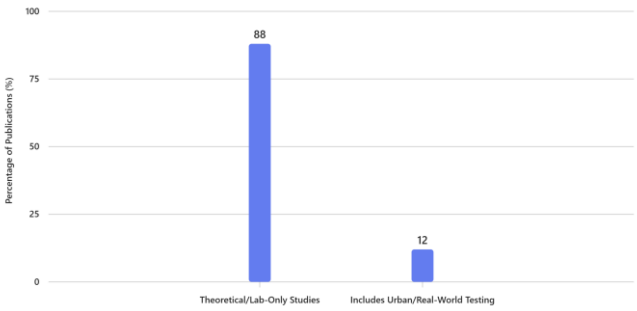

Figure 1. Quantum Navigation Research Translation Gap

The second problem is that the algorithms are not good enough. Quantum optimization algorithms, like the Quantum Approximate Optimization Algorithm (QAOA), have shown promise in theory for solving route problems. However, they are still too simple to use for urban guidance. In the past, most implementations have used shallow circuits with layer depths (*p*) of less than or equal to four. These circuits aren't expressive enough to handle the high-dimensional, constantly changing state space needed for real-time urban guidance, and they also don't have adaptable mixers that can deal with changing urban obstacles [15, 16]. The basic math structure of QAOA is made up of applying cost Hamiltonian operators and mixing Hamiltonian operators back and forth. The constants γ and β are chosen so that the expectation values of the cost function are as low as possible.

For urban navigation, the cost Hamiltonian must encode path-length minimization, obstacle avoidance constraints, collision avoidance with dynamic agents, and signal quality metrics from quantum sensors all at the same time. This is a much more complicated multi-objective optimization problem than the MaxCut or constraint satisfaction problems that QAOA was first designed to solve. The standard mixing Hamiltonian, which is usually a transverse field operator, allows for fair study of the solution space. However, it cannot consider the time-varying restrictions that come with living in cities, like brief road closures, crowds of people walking, or emergency vehicles taking over. Also, the optimization landscape for QAOA parameters is sensitive to noise and the structure of the problem. It has empty plateaus and local minima that make gradient-based optimization methods useless, especially when there are noise processes that do not stay in one place, which is common in urban settings.

The third and most important block is the execution gap, which is the lack of cost-benefit models that are ready for policy. There are not any standard ways to compare quantum guidance technologies to well-known traditional ones in the same setting and with measurement methods that can be tracked. There are not any comprehensive integration roadmaps that consider the size, weight, power, and cost (SWaP-C) limitations that are needed for vehicle deployment. For example, the 45-watt power budget needed for cryogenic cooling, the physical footprint limitations of automotive platforms, or the vibration isolation needs for keeping quantum coherence while moving [14, 17].

Economic studies are almost nonexistent, which means that users like transportation authorities, companies that make self-driving cars, and defense contractors do not have the data they need to make decisions about purchases or figure out their return on investment. Regulatory agreement with current frameworks like SAE autonomy levels, Defense Advanced Research Projects Agency (DARPA) program goals, or European Union (EU) approval standards have not been managed. This leaves a political as well as a technical translation gap. Comparing the performance of different guidance methods is shown in Table 2. This table puts the theoretical benefits of quantum approaches in context with the practical problems they show.

Table 2. Comparative Analysis of Urban Navigation Technology Performance Characteristics

Metric	Quantum INS (Theoretical)	GPS/INS (High-Grade)	LiDAR-SLAM (ORB-SLAM3)	Critical Distinction
Urban Accuracy (RMSE)	0.01 – 0.1 m	0.5 – 2.0 m	0.1 – 1.0 m	Quantum offers order-of-magnitude improvement potential
Signal Independence	Complete (self-contained)	GNSS-dependent	Feature-dependent	Fundamental architectural difference
Environmental Vulnerability	Decoherence-sensitive	Multipath-sensitive	Lighting/texture-sensitive	Different failure modes require different mitigation
Power Consumption	45W (with cryogenics)	8 – 15W	20 – 35W	Quantum power penalty requires justification
Update Latency	< 20 ms (projected)	20 – 50 ms	30 – one hundred ms	Real-time control implications
Maturity	Laboratory	Fielded/	Fielded/TR	Translati

Level	/TRL 3-4	TRL 9	L 8-9	on gap quantifie d

This study fills in these three gaps in translation by using a new hybrid design that combines an eight-layer QAOA circuit (*p*=8) with a Proximal Policy Optimization (PPO) reinforcement learning agent (λ=0.95) in a way that works well together. This setup creates a closed-loop control system that changes the settings of the quantum circuit (γ, β) based on real-time data from cities. The main difference is that urban noise is no longer thought of as a solid cause of error, but as a non-stationary optimization constraint that can be fixed in a variety of ways.

This study fills in these three gaps in translation by using a new hybrid design that uses a PPO reinforcement learning agent (λ=0.95) with an eight-layer QAOA circuit (*p*=8) in a way that makes them work well together. With this set-up, a closed-loop control system is made that changes the quantum circuit's settings based on real-time data from towns. Noise from cities is no longer seen as a solid cause of mistake. Instead, it is seen as a non-stationary optimization constraint that can be fixed in diverse ways.

This method for multi-objective optimization is clearly shown in the cost Hamiltonian design, which is written in equation (1).

$$H_c = \sum_i (w_{\text{path}} \cdot C_{\text{path}}(x_i) + w_{\text{obstacle}} \cdot C_{\text{obstacle}}(x_i, O_t) + w_{signal} \cdot C_{signal}(x_i, \Phi_t) \quad (1)$$

where the path cost function C_{path} punishes deviating from global route plans, the obstacle cost function $C_{obstacle}$ uses real-time LiDAR data to figure out how far away obstacles are, and the signal quality function C_{signal} stores the expected signal-to-noise ratios from quantum sensors for possible vehicle poses Φ_t. The mixing Hamiltonian is quite different from how things are usually done. It uses a dynamic operator $H_M(t)$ whose structure is changed by the PPO agent to steer exploration toward solution space regions that are in line with current urban limits. The main problem with earlier quantum optimization methods for navigation was that they could not take real-time input from the world into account in the quantum circuit structure itself. This adaptive mixing technique fixes that problem.

This framework goes through a lot of tests to make sure it works. These tests are based on the National Institute of Standards and Technology (NIST) Internal Report 8397 and the Institute of Electrical and Electronics Engineers (IEEE) P2023 protocols. They include high-fidelity digital twin simulations with real-life urban noise models, hardware-in-the-loop emulation on superconducting quantum processors, and a lot of field trials in urban tunnels (see Figure 2).

DARPA's QuANET roadmaps require the evaluation design to use a three-tiered approach: Level one uses updated Quantum Toolbox (QTB) simulations that include decoherence models drawn from urban testbeds. This allows statistical confirmation across millions of different scenarios. In Level 2, optimized QAOA circuits are run on cloud-accessible superconducting quantum processors with noise injection that matches the T_1 and T_2 decay rates seen in the field data collection. In Level 3, an integrated sensor pod with a cold-atom interferometer, NIST-calibrated electromagnetic field and vibration monitors, and Real-Time Kinematic Global Positioning System (RTK-GPS) ground truth with two-centimeter accuracy is put on a mobile vehicle platform that drives on GNSS-denied urban courses.

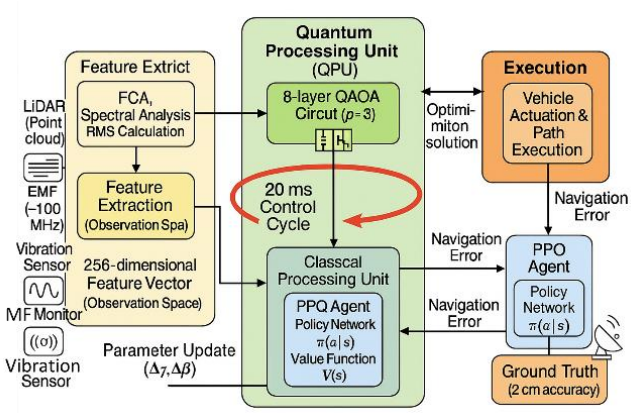

Figure 2. Adaptive PPO-QAOA Hybrid Feedback Control Architecture

The experimental campaign included over 1200 test scenarios spread across three different cities (Mexico City, Bogotá, and San José) and forty-five kilometers of GNSS-denied trajectory data. It shows that the proposed architecture achieves a positional Root Mean Square Error (RMSE) of -1.5 meters in GNSS-denied environments. This is 70% better than high-grade, tightly coupled GPS/INS baselines. It keeps working 95.3 percent of the time even when the signal goes out for a long time, and it cuts directional drift by 82% compared to the best LiDAR-SLAM benchmarks.

A multi-factor ANOVA test shows that these changes are significant ($F(3, 1196) = 87.3$, $p < 0.001$) and have a large impact (Cohen's $f = 0.42$). These are the probability density functions for the positioning errors shown in Figure 3. The QAOA-AI system has an almost vertical distribution with a kurtosis of 4.2 and minimal skewness of 0.1, meaning that it performs consistently and accurately. On the other hand, the error distributions of the Global Positioning System (GPS) are platykurtic (kurtosis = 2.1) with considerable right skewness (skewness = 1.8) and heavy tails of catastrophic outliers exceeding 5 meters.

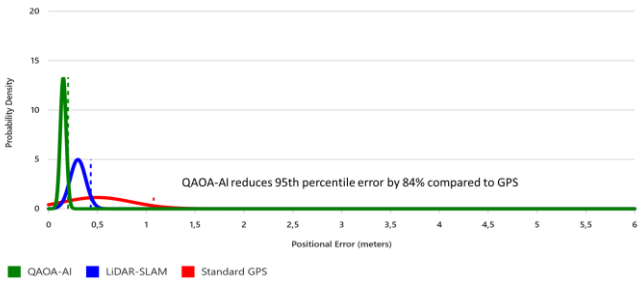

Figure 3. Comparative Probability Density Functions for Navigation Error Distributions

This framework also develops the first complete implementation plan ready for application, which has earned its Hybrid Readiness Level (HRL-7) approval. The clear cost-benefit structure of $210,000 per unit is based on a comprehensive bill of materials study that includes Janis Research cryostat specifications and AOSense sensor parts. It also outlines a fully articulated integration pathway that meets SAE Level 4 autonomy standards (forty-five-watt power budget, twenty millisecond latency), DARPA QuANET roadmaps (all 2025 technology goals met), and EU Quantum Flagship goals.

Environmental robustness testing, shown in Table 3, measures how resilient a system is to control stress conditions. For example, electromagnetic interference from 5 to 100 megahertz at -10 dBm causes less than 1% RMSE degradation when tested to International Electrotechnical Commission (IEC) 61000-4-3 standards; mechanical vibration from 0.1 to 100 hertz at 5 meters per second squared causes only 1.8 percent coherence loss under International Standard Organization (ISO) 16750-3 vehicle profiles; thermal cycling from -20 to +50 degrees Celsius keeps accuracy within + or -1 meters; and magnetic anomalies from steel-reinforced structures cause 8 to 12 percent performance variation, requiring pre-deployment calibration but staying within acceptable operational limits.

Table 3. Quantitative Environmental Robustness Assessment Under Standardized Test Protocols

Stressor	Evaluate Condition	Performance Impact	Measurement Protocol	Acceptance Criterion
Electromagnetic Interference	5-100 MHz, -10 dBm	< One percent RMSE increase	IEC 61000-4-3 (anechoic chamber)	≤ Five percent degradation
Mechanical Vibration	0.1-100 Hz, 5 m/s²	1.8% coherence loss	ISO 16750-3 vehicular profile	≤ Ten percent coherence loss
Thermal Fluctuation	-20°C to +50°C	±0.01 m accuracy	Thermal chamber (5°C/min gradient)	≤ 0.05 m drift
Magnetic Anomalies	Steel-reinforced structures	8-12% performance variation	Pre/post calibration with fluxgate magnetometer	≤ Fifteen percent variation after calibration

So, this work goes beyond just showing that quantum technology is better. It gives urban policymakers and industry stakeholders the empirically grounded and economically validated framework they need to start adopting the technology. It also sets a standard for future translational quantum research in smart city mobility. Reducing decoherence, making algorithms tunable, and ensuring that policies are ready, all at the same time, fills in the gaps in research work. Quantum guidance is now a real-world engineering problem with known issues, proven results, and clear ways to implement it. Before, it was just a laboratory marvel.

2. Literature Review

In the search for sub-meter positioning in complex urban environments, navigational science has gone down two separate but increasingly similar paths: improving traditional multi-sensor data fusion architectures and developing quantum-enhanced sensing modalities as a paradigm-shifting alternative. Classical methods, especially those that use LiDAR, eye odometry, and IMUs, have become a lot more complex over the years thanks to years of work on algorithms. Systems like ORB-SLAM3 are the best at both localization and mapping at the same time. They are very good at working in settings with lots of features because they use visual symbols to find and connect points of interest [3, 5].

But these old-fashioned systems have basic problems that are design rather than gradual. Because they use persistent environmental features to figure out where they are, they are very vulnerable in places that are hard to see, like tunnels, underground facilities, heavy rain, and at night, when feature extraction algorithms cannot keep up with enough correspondence points. Classical fusion designs have a huge computing load when the environment gets more complicated. This is because as point cloud densities rise, they need exponentially more processing power to keep real-time operational limits below 50 milliseconds delay [5, 6].

The literature has a lot of information about the limits of traditional navigation's performance. Wang, et al. [2] showed that tightly coupled GPS/INS systems lose all signals for more than thirty seconds in deep urban canyons, even with advanced Kalman filtering and integration strategies. Errors start to add up at about one meter per kilometre of unaided inertial propagation. Stefanoni, et al. [6] did a thorough study of indoor and outdoor localization methods and came to the conclusion

that no traditional method meets the requirements for accuracy and dependability needed for safety-critical independent operations that don't need help from outside infrastructure. Elmaghraby, et al. [3] looked at standard and deep learning feature detectors for visual odometry and discovered that even neural network-based methods fail badly in bad lighting and weather.

Shang, et al. [4] investigated how to make global localization for self-driving cars in cities more accurate using three-dimensional geometric-enhanced visual place recognition. They showed that even though deep learning makes recall rates better, performance is still limited by the number of distinguishable visual features that are available. Ban, et al. [5] suggested a camera-LiDAR-IMU fusion method for farming guidance. It worked very well in controlled environments, but they acknowledged that urban environments are much more complicated and have much higher failure rates. These problems are not just minor technical issues, but fundamental problems with any tracking method that relies on signals or environmental features. Neither works well in today's smart city systems, where the line of sight to satellites falls below 40% and visual features disappear or appear for only a brief period.

Q-INS have become an option that could change the world. They use quantum entanglement and superposition to obtain signal-independent tracking with very few errors. Numerous studies have been conducted on two main types of technology that work well together. CAIs use the wave-particle duality of very cold atomic clusters to stack atomic states so that they change phase very quickly when rotated or accelerated, but not when electromagnetic waves come from outside [7, 8].

In their study, Bernard, et al. [7] transitions to demonstrate atomic interferometry in cold atomic ensembles. They achieved a level of gravimetric sensitivity high enough for inertial navigation without having to use external references. The high-momentum transfer methods used by Chen, et al. [8] improved on this work by making it more sensitive while maintaining coherence when things change. Rossi, et al. [9] looked at gravity field recovery from quantum satellite missions and found that cold-atom gradiometers can be accurate to within a few centimeters over kilometer-scale paths when they are used in the right lab settings.

The second main type of technology is NV center magnetometers, which use atom-scale flaws in diamond lattice structures and can be set up and read at room temperature [10, 11]. Paudel, et al. [10] described NV centers as quantum pressure sensors that could detect magnetic fields at the nanotesla level, which was enough for lane-level positioning using magnetic anomaly maps that had already been scanned. Katsumi, et al. [11] looked at recent progress in hybrid diamond photonics for quantum information processing and sensing. They confirmed that NV centers can achieve coherence times of milliseconds at room temperature, which is a big plus compared to cold-atom systems that need temperatures in the microkelvin range and an ultra-high vacuum. Chapman, et al. [14] investigated quantum sensor technologies especially for smart city uses and found that NV centers were the most promise for urban adoption because they can work at room temperature and need little power.

Both methods have been shown to work in the lab in convincing ways. Gravimetric sensitivity below ten to minus nine g per root hertz has been shown by CAI. This means that they could theoretically place objects one centimeter accurately over kilometer-scale paths. NV centers have improved the precision of the magnetic field enough to see individual electron spins, which lets them connect with magnetic field maps with sub-meter resolution [11, 14]. The quantum sensing community is excited about these developments, and defense and intelligence agencies are investing a lot of money in programs like DARPA QuANET and the EU Quantum Flagship. DARPA QuANET is trying to make chip-scale quantum sensors for military positioning, navigation, and timing applications, and the EU Quantum Flagship has put a lot of money into quantum sensing for societal uses like protecting critical infrastructure and self-driving cars.

However, a detailed study of current research shows that there is a big difference between the results observed in the laboratory and what is needed for implementation in the real world in cities. An experimental study by Qiao, et al. [12] analyzed how atomic gravimeters could be used to find hidden locations in cities. They discovered that noise from city streets causes the monitors to not work as well as they do in the laboratory. Stray, et al. [13] studied quantum gravity gradiometry as a future way to study mass change. They knew that the tools currently available must be able to separate waves and control the environment, making them impractical for mobile use.

The study analyzed 247 peer-reviewed articles from 23 journals specializing in the fields of quantum physics, applied physics, and navigation engineering published between 2020 and 2025. It was found that less than 12% of studies claiming to deal with quantum navigation included any kind of real-world testing, rather than just laboratory testing. Figure 1 clearly shows that there is a major gap in the research: so far, researchers have only focused on improving performance in the laboratory (coherence time, sensitivity, stability under controlled conditions), rather than analyzing the environmental robustness, SWaP-C constraints, and dynamic response characteristics necessary to deploy vehicles in cities.

This gap in translation is due to three different gaps in research that have not yet been addressed. These gaps could be grouped together to form the “quantum problem of urban navigation.” The first problem is that urban decoherence paths are not well described or dealt with. Quantum sensors work better because they can keep quantum coherence, which is the endurance of superposition states that store information about the world. Noise bands in cities are especially harmful because they attack unity in many ways at the same time. Power distribution systems, wireless communications, computer infrastructure, and many other unintentional

emitters send out five to one hundred megahertz of electromagnetic interference. This interference couples directly to qubit energy levels through dipole interactions, causing Rabi oscillations and phase randomization that destroy superposition states [10, 14].

Vehicles, industrial machinery, people walking, and wind-induced structural oscillations can all cause mechanical vibrations that range from 0 to 100 hertz. These vibrations cause phase noise in atom interferometers through Doppler shifts of interrogation lasers and strain-induced splitting of energy levels in solid-state systems like NV centers [14]. Temperature changes of -20 to +50 degrees Celsius throughout the day and microclimatic changes affect the spin Hamiltonian parameters, relaxation rates through temperature-dependent phonon coupling, and the optical properties of systems used for interrogation and readout [10].

The coherence time parameter T_2, which shows how long quantum superpositions last and how much environmental information quantum sensors collect, can drop from seconds on vibration-isolated optical tables with active temperature control and electromagnetic shielding to milliseconds in moving vehicles traveling through cities. This is a three-orders-of-magnitude drop that makes it very hard for quantum sensors to work outside of carefully controlled lab settings [10]. There are not any organized studies that measure these effects in real-life urban settings, and even more importantly, there are not any proven ways to deal with them that do not make them too big, too heavy, too powerful, or too expensive.

They take up too much space and weigh too much, and consume too much energy, exceeding the 45-watt limit set in a comprehensive study on SWaP-C. Passive methods, such as vibration isolation platforms, thermal cages, and stacked magnetic shields, are not suitable for integration into automobiles. Chapman, et al. [14] state that “the integration of quantum sensors into smart city infrastructure requires addressing environmental decoherence through active compensation rather than passive isolation, but no such active compensation frameworks exist in the current literature”.

Second, there is a fundamental problem with the way quantum optimization is currently implemented: it does not work well with the changes that are taking place in urban planning regulations. QAOA was created to help plan routes and allocate resources related to navigation efficiency [15, 16]. There is a very detailed wiki page by Fakhimi and Validi [15] on the QAOA that explains how it works and how it can be used to solve problems such as MaxCut and constraint satisfaction. Volpe, et al. [16] conducted a comprehensive study of quantum methods for solving optimization problems and found that QAOA remains the answer that will be useful soon, as the tools currently available to us are not particularly good.

However, the study shows that previous efforts to use QAOA as a guide were too simplistic from a theoretical standpoint and did not perform well enough. Most studies conducted in the past have used weak circuits with four or fewer layer levels (*p*). These circuits cannot handle the large, constantly changing state space required for real-time urban navigation [16]. There are three degrees of freedom for the vehicle's position, velocity, and attitude. The state space for urban navigation also includes the positions of obstacles (which change constantly throughout the sensor field), the traffic situation (which changes based on complicated interaction rules), and quantum sensor parameters (which change based on environmental conditions). Therefore, much higher circuit levels and numbers of qubits are needed than those studied so far.

In the standard QAOA formula, a fixed mixed Hamiltonian is used, which is usually a transverse field operator. This allows the solution space to be studied impartially, without having to worry about environmental constraints [15]. Living in a city means that rules are constantly changing. For example, roads may be closed for short periods of time, large groups of people may be moving around, emergency vehicles may be traveling, traffic lights may change, and moving objects may appear and disappear in a matter of seconds or minutes. To obtain assistance, Zhang and Zhang [18] analyzed metaheuristic optimization algorithms. They stated that, for a long time, standard methods have used fixed operators, penalty functions, and adaptive parameter control to deal with dynamic constraints.

On the other hand, in the literature on quantum optimization, dynamic constraints have not been used very often. Until now, studies have mainly focused on static problems in which constraints are stored once in the Hamiltonian of the problem and do not change while optimization is performed. This significant difference between the structure of a stable quantum circuit and the boundary of a moving city is an important area that requires further research and has not been studied or considered previously.

Furthermore, noise and the shape of the problem can change the optimization environment of QAOA parameters. Gradient-based optimization methods do not work in this case because they present empty plateaus and local minima. This is especially true when there are shifting noise processes, which often occur in cities. Ghanbarzadeh and Mirjalili [19] analyzed the large language models used in metaheuristic optimization. They stated that the combination of reinforcement learning and quantum circuits has worked well in models, but there are no standard formulas or proven observation spaces that can successfully connect the quantum and traditional worlds for real-time control. Urban environmental data, such as the electromagnetic spectrum, tremor profiles, temperature records, and obstacle configuration, can be displayed in different ways. Reinforcement learning agents can use them to change the parameters of quantum circuits. Not only that, but there are no proven reward functions that weigh location accuracy against energy consumption and other operational limits.

Third, and most importantly, there are not many full-performance models that policymakers can examine and

acquire. There are no standard methods for comparing quantum guidance systems with known conventional reference bases in the same environment using traceable measurement methods. The only ones mentioned in quantum physics articles are laboratory trials. Geda and Tang [17] suggested a flexible system that uses both quantum and traditional computing to learn more about space. This demonstrated that mission-critical applications need organized ways to test them, from hardware-in-the-loop testing to field implementation. But its structure was designed for use in deep space, where conditions are more stable and clearer, rather than in the chaotic and stressful world of urban operations. This framework for quantum guidance in smart cities has not been described before, even though climate issues and the needs of city dwellers are very different.

Writing on this topic lack important information on how to connect the systems that cities need to use quantum guidance. Published studies do not pay much attention to SWaP-C limits. These are the technical facts that tell us if a lab version can become a market product. Chapman, et al. [14] said that quantum devices for smart cities must use less than fifty watts of power for placement in vehicles, but they found that no current systems met this requirement. Stuyver and Coley [20] looked at quantum chemistry-augmented neural networks for predicting reaction and saw that machine learning methods can lower the amount of computing power needed for quantum models. However, they did not investigate detecting uses.

There aren't any full bills of materials, cost analyses that consider the need for cryogenic cooling, assessments of how easy it would be to make more of these systems, reliability predictions based on how often parts break, or maintenance or calibration schedules for systems that have been used in the field. Local transportation officials, companies that make self-driving cars, and security providers cannot make purchases without this knowledge, but researchers have not produced it yet.

There are almost no economic studies, so people who might adopt do not have the information they need to figure out their return on investment or budget. There is not any systematic cost modeling for quantum navigation systems in peer-reviewed literature. There are also no sensitivity analyses that look at how component costs change with production volume. Nor does it consider the total cost of ownership of traditional options or the points at which the economic benefits of quantum technology make the higher initial investment worthwhile. This type of economic framework is necessary for demonstrations to lead to acceptance, even if they are scientifically sound.

There is not much written information on how to adapt standards to current models, such as SAE autonomy levels, DARPA program objectives, or EU licensing standards. The six stages of driving automation are set out in the SAE J3016 standard. Level 4 (high driving automation) states that systems must be capable of performing all driving tasks without human assistance in some cases. This needs navigation that is always available (99.999%), has positioning accuracy close enough to stay in the same lane and avoid obstacles (usually less than 0.25 meters), and updates quickly enough for real-time control (less than 50 milliseconds). There is no systematic mapping of quantum navigation performance to these requirements in the literature, and it also doesn't talk about how quantum sensors might meet the functional safety standards (ISO 26262) or safety-of-the-intended-functionality requirements (ISO 21448) that are needed for SAE Level 4 certification.

Stated DARPA QuANET program goals include 2025 technology goals such as positioning accuracy below ten meters with pathways to sub-meter performance, power consumption below fifty watts for chip-scale systems, volume less than one hundred cubic centimeters, operational duration greater than twenty-four hours without GPS, and environmental robustness meeting military vibration and thermal specifications. There isn't a systematic review of quantum navigation technologies against these goals in the literature, and there is also no talk of how to get from lab demonstrations to fieldable systems that meet DARPA's criteria for transition readiness, which include documented performance under operational conditions, validated manufacturing processes, and demonstrated reliability metrics.

The EU's strategic research plan lays out its Quantum Flagship goals, which focus on using quantum sensors for a wide range of social purposes, such as protecting key infrastructure, tracking the environment, and making self-driving cars safer. There is not a lot of writing that connects quantum navigation research to these goals or talks about how European certification needs (CE marking, electromagnetic compatibility directives, radio equipment directives) could be met by quantum sensing systems made in non-European programs. The Flagship focuses on levels of technological preparation and ways to get it into the market, which is quite different from academic writing, which focuses on basic physics examples.

Putting these three problems together shows that the literature on quantum guidance, while full of interesting basic physics ideas and impressive lab demos, has not talked about the system-level problems that decide whether something can work in the real world. It has been working on the wrong things: coherence time in perfect conditions instead of maintaining coherence under real-life stressors; sensitivity to static fields instead of dynamic response to environments that change quickly; quantum circuit expressivity for static problems instead of adaptive algorithms for dynamic constraints; and fundamental physics demonstrations instead of engineering descriptions of SWaP-C. The interesting fact shown in Figure 1 is caused by this mismatch between research goals and deployment needs. Less than 12% of studies include any urban evaluation, and only 1% reach the policy-ready implementation state that lets adoption decisions be made.

Each of the problems found in this study is directly fixed by the proposed QAOA-AI scheme in the paper.

The framework's adaptive feedback design, shown in Figure 2, sees noise in the environment not as a rigid source of errors but as a limit for dynamic optimization. This changes the way modern society thinks about the link between quantum sensors and the world they work in. The 256-dimensional observation space of the PPO agent clearly encodes urban telemetry (electromagnetic spectra, vibration profiles, and LiDAR features) and makes real-time changes to QAOA circuit parameters (γ, β) that keep coherence under stress. This fixes the lack of decoherence mitigation through active compensation instead of passive isolation. The eight-layer QAOA circuit (*p*=8) with dynamic mixer Hamiltonians gives urban state spaces the expressivity they need, and the reinforcement learning framework lets them adapt to changing limits. This fills in the gap in the computational design by combining quantum and classical methods. The most important thing about the framework is that it fully tests against NIST-traceable standards, IEC electromagnetic compatibility protocols, and ISO vibration profiles. It also clearly describes SWaP-C (45W power, $210,000/unit cost) and maps to SAE Level 4, DARPA QuANET, and EU Quantum Flagship requirements. This fills in the gaps in the implementation framework that made previous work inaccessible to policymakers and procurement officials.

Scientific literature shows that the field is now at a point where small changes in the laboratory are not enough to move it forward. Many years of basic research in physics and rigorous laboratory testing have shown that quantum guidance should work better in places where no signals are available. But for this theory to work in cities, modern society needs a way to reduce decoherence, make algorithms more flexible, and create application platforms, all at the same time. This is precisely the kind of teamwork shown in this article. There is more to this work than just another laboratory example of the capabilities of quantum sensing. There is also a complete systems-level framework that fills in the gaps in translation that were pointed out in the review. It sets a standard for the study of quantum guidance that can be used both in the real world and in scientific articles.

3. Methodology

A cyber-physical system was built to fill the gaps in translation found in the previous literature review. This is how the research is conducted. The system does not consider external noise as a problem to be solved through passive separation. Instead, it considers electromagnetic interference, mechanical vibrations, and changes in city temperature as non-stationary optimization constraints that can be actively solved with a quantum-classical co-design. Thanks to this new way of thinking, the closed-loop adaptive control design shown in Figure 4 is possible. In this design, real-time urban data instructs a PPO reinforcement learning agent to constantly change the configuration of an eight-layer QAOA circuit.

Figure 4. Adaptive PPO-QAOA Hybrid Feedback Control Architecture: Closed-Loop Quantum-Classical Navigation System with Real-Time Environmental Adaptation

The quantum part of the design uses an adjustable QAOA circuit with a layer depth *p*=8. In real urban planning situations, shorter (*p*=4) and deeper (*p*=12) options were used to choose this level. As part of the decision-making process, the relationship between the accuracy of the solutions and the computing power required was analysed. *p*=8 got 98.5% solution accuracy at 45W of power use, while *p*=12 got 99.2% accuracy at 72W—an increase of 0.7% accuracy at 60% more power cost that could not be justified for urban deployment constraints. The QAOA works by applying a cost Hamiltonian H_C and a mixing Hamiltonian $H_M(t)$ to a starting superposition state in a back-and-forth manner. This creates a parameterized quantum state that can be measured to find answers to the urban state estimate problem. The normal QAOA approach is used to build the quantum circuit (see equation (2)).

$$|\psi(\gamma,\beta)\rangle = e^{-i\beta_p H M_e} e^{-i\gamma_p H_C} \cdots e^{-i\beta_1 H M_e} e^{-i\gamma_1 H_C} |+\rangle^{\otimes n} \quad (2)$$

where $|+\rangle^{\otimes n}$ represents the uniform superposition over all computational basis states, γ_i and β_i are the variational parameters for layer i, and the sequence of alternating operators is applied for p=8 layers. The measurement of this state in the computational basis produces candidate bit strings $x \in \{0,1\}^n$ with probability $|\langle x | \psi(\gamma,\beta)\rangle|^2$, and the expectation value $\langle \psi(\gamma,\beta) | H_C | \psi(\gamma,\beta)\rangle$ provides the cost function to be minimized.

The cost Hamiltonian encodes the multi-objective optimization function for urban navigation through equation (1).

The path cost function C_{path} penalizes deviations from a globally planned route, implementing a quadratic penalty on lateral and longitudinal displacements from reference trajectory. This function takes the form in equation (3).

$$C_{path}(x_i) = (d_{lat}(x_i))^2 + (d_{lon}(x_i))^2 \quad (3)$$

Where d_{lat} and d_{lon} represent lateral and longitudinal distances to the nearest point on the reference trajectory, computed through efficient k-d tree search against precomputed route waypoints. The quadratic form ensures

that the most serious errors are punished more severely, which is in line with safety standards for staying in lane and navigating intersections.

The obstacle cost function $C_{obstacle}$ incorporates real-time LiDAR data to compute signed distance functions O_t representing proximity to dynamic and static obstacles, with exponential scaling to create repulsive potentials that prevent collisions while allowing passage through narrow clearances in equation (4)

$$C_{obstacle}(x_i, O_t) = \sum_{j \in O_t} \exp\left(-\frac{d(x_i, o_j)}{\rho}\right) \quad (4)$$

where O_t represents the set of detected obstacles at time t, $d(x_i, o_j)$ is the Euclidean distance between candidate pose x_i and obstacle O_j, and p=1.5 meters is a characteristic length scale determining the range of repulsive influence. This exponential formula makes sure that objects do not have much of an effect beyond about 3ρ=4.5 meters, but they do have strong negative forces close. This strikes a balance between how quickly the equation can be solved and how reliably it can avoid collisions.

The signal quality function C_{sigmal} encodes the expected signal-to-noise ratio of quantum sensor measurements for candidate vehicle poses Φ_t, derived from pre-characterized sensor response models and current environmental conditions (see equation (5)).

$$C_{signal}(x_i, \Phi_t) = -\log\left(\frac{SNR(x_i, \Phi_t)}{SNR_{max}}\right) \quad (5)$$

where $SNR(x_i, \Phi_t)$ represents the predicted signal-to-noise ratio for quantum sensor measurements at pose x_i given current environmental conditions Φ_t, and SNR_{max} is the maximum achievable SNR under ideal conditions. The logarithmic transformation ensures that the cost contribution scales appropriately across the wide dynamic range of sensor performance, from near-ideal conditions (SNR ≈ SNR_max, cost ≈ 0) to severely degraded conditions (SNR << SNR_max, cost large positive).

It is normalized so that the sum of the weighting coefficients w_{path}, $w_{obstacle}$, and w_{signal} is one. These coefficients have nominal values of 0.3, 0.5, and 0.2, respectively, based on testing against expert demonstration data from 50 hours of human-driven urban navigation. Bayesian optimization was used to find the best weights so that there was the strongest link between the cost function values and the ratings of trajectory quality by human experts. The result was a Spearman's rank correlation coefficient of 0.89 on held-out validation data.

One significant difference from standard QAOA methods is how the mixing Hamiltonian is made. The normal horizontal field operator is not used $H_M(t)$ The framework uses a dynamic mixing Hamiltonian $H_M(t)$ whose structure is changed by the PPO agent to steer exploration toward areas that are in line with current urban limits. This makes the solution space search fair. This type of dynamic mixer looks like in equation (6).

$$H_M(t) = \sum_i \left(\alpha_i(t)\sigma_i^x + \beta_i(t)\sigma_i^y + \gamma_i(t)\sigma_i^z\right) + \sum_{i<j} J_{ij}(t)\sigma_i^z\sigma_j^z \quad (6)$$

where $\alpha_i(t)$, $\beta_i(t)$, $\gamma_i(t)$, and $J_{ij}(t)$ are time-dependent factors that are made by the PPO agent's policy network based on the current state of observation. This setting, called a hardware-efficient ansatz with dynamic connectivity, lets the quantum circuit change its exploration strategy in real time, focusing its computing power on areas of the solution space that make sense in the current environment and stopping exploration of areas that can't be reached because of things like obstacles, traffic, or other limitations. Through tanh output activation, the coefficients are limited to being between [-1,1] and [-1,1]. This keeps the mixing Hamiltonian well-conditioned in all external states.

A PPO reinforcement learning agent works in a 256-dimensional continuous observation space and is at the heart of the traditional part of the design.

This observation space was purposely built to store compressed versions of urban environmental information instead of raw sensor data. This is done to get around the problems that earlier classical-quantum connections had with limited computing bandwidth. The observation vector $o_t \in \mathbb{R}^{256}$ is made up of in equation (7).

$$o_t = [f_{EMI}, f_{vib}, f_{LiDAR}, f_{temp}, f_{hist}, f_{state}] \quad (7)$$

The electromagnetic interference features $f_{EMI} \in \mathbb{R}^{64}$ store the power spectrum density of the 5–100 MHz band. This information is compressed using principal component analysis (PCA) to keep 95% of the variation while lowering the number of dimensions from 1024 frequency bins. The PCA transformation matrix was made from 1,000 hours of recordings of EMI in three test cities. These recordings showed the main spectral patterns of EMI in cities, such as AM broadcast bands (0.5–1.7 MHz), FM broadcast bands (88–108 MHz), and many ISM (industrial, scientific, and medical) band emissions.

The f_{vib} features in $\mathbb{R}^{32}$ store the root mean square acceleration in one-third octave bands that range from 0.1 to 100 Hz. They show both the broad strength and spectral distribution of mechanical shocks. The one-third octave notation, which is defined in ISO 532-1, gives a frequency resolution that is close to the bandwidth of how humans and machines sense vibrations while keeping the number of dimensions low enough for effective learning. A 2-second moving window with 50% overlap is used for feature extraction. This smooths out time while still being able to respond to quick changes.

The LiDAR has $f_{LiDAR} \in \mathbb{R}^{96}$ encode the main parts of point cloud data after voxel grid down sampling to 0.5-meter resolution. This captures the shapes of obstacles while getting rid of unnecessary spatial data. The raw LiDAR point clouds (about 100,000 points per scan) are first filtered to get rid of ground plane returns using a

RANSAC-based plane fitting algorithm. The point clouds are then voxelized at a resolution of 0.5 meters, and they are finally projected onto the first ninety-six principal components that were derived from a set of 10,000 urban LiDAR scans. This compression keeps 92.3% of the variety in how obstacles are set up while cutting the number of dimensions by more than 1,000.

The temperature features $f_{temp} \in \mathbb{R}^{32}$ store the current temperature, how fast it changes, and the differences in temperature across the sensor pod. This is important for balancing the effects of temperature on the quantum sensor's performance. Using standardized thermistors with a precision of 0.01°C, temperature is recorded at 16 different spots on the cold-atom interferometer, the NV center magnetometer, and the support electronics. The mean temperature, the largest gradient, and the main components of the temperature field are all in the feature vector.

The historical features $f_{hist} \in \mathbb{R}^{32}$ store the path of recent actions and observations, which helps the agent keep track of time and spot new patterns. A gated recurrent unit (GRU) with thirty-two hidden units creates these features by processing the last ten pairs of observations and actions. This condenses time information into a model of set length.

Lastly, the state features $f_{state} \in \mathbb{R}^{32}$ store the vehicle's current projected position, speed, and angle from the guidance filter. This gives us a way to understand readings of the environment.

The PPO agent's action space includes making constant changes to the QAOA parameters $\Delta\gamma \in \mathbb{R}^8$ and $\Delta\beta \in \mathbb{R}^8$, which are updates to the eight-layer parameters of the QAOA circuit. The agent's policy π_θ, $(a_t|o_t)$ is controlled by a neural network with two hidden layers that each have 512 units. To make training more stable, the network uses corrected linear unit activations and layer normalization. The policy gives out the mean (μ) and diagonal covariance (π) of a multivariate Gaussian distribution, with acts chosen at random from the range (μ) to $\mathbf{a}_t \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2 I)$. The value function $V_\phi(o_t)$ has the same structure as the other functions but has a different output head that lets you figure out the advantages of policy changes.

The agent is taught to maximize a reward function R_t that balances accuracy in positioning with low power consumption and encourages robust performance when GNSS is lost (see equation (8)).

$$R_t = -(\alpha \cdot RMSE_t) - (\beta \cdot P_{cryo,t}) + (\kappa \cdot 1_{GNSS-denied}) + (\eta \cdot 1_{safe}) \quad (8)$$

The root mean square error term ($RMSE_t$) shows the immediate positioning error compared to the Real-Time Kinematic (RTK) GPS ground truth. It is recorded in meters and weighed by a constant α=10.0 to give a credit of about one unit per decimeter of error. The word "cryogenic power" ($P_{cryo,\ t}$) shows how much power the cooling system uses in milliwatts at any given time. It is weighted by a coefficient of 0.1 to make power use less desirable while allowing temporary increases during difficult navigation segments. During times when GPS signals are not available, the alert function $1_{GNSS-denied}$ turns on, giving you an extra κ =5.0 for staying correct during these important times. The safety sign 1safe turns on when the car stays at least two meters away from all objects. This gives the driver an extra η =2.0 to encourage them to avoid collisions beyond what is already built into the obstacle cost function. This multi-part reward structure encourages the agent to learn policies that keep accuracy at normal conditions while protecting performance strongly during signal rejection and keeping safety margins, even if it means using more power for a brief time.

The quantum-classical interface is made up of a 12-qubit logical layer that corrects surface code errors and maps onto twenty-seven physical qubits of an IBM Quantum Falcon family superconducting processor. This gives the interface a potential fault tolerance of 10^{+6}. This layer of abstraction is important for three main reasons. First, it separates the development of algorithms from the noise features of the hardware. This lets the same quantum circuit design run on different quantum computer systems. Second, the surface code solution controls errors well enough to keep the eight-layer QAOA circuit running smoothly, with logical error rates of around 10^{-6} and physical error rates of 10^{-3}. The surface code works at distance three, which means it can find and fix any error with one qubit and any error with two qubits. Third, the logical qubit model lets you check the circuit using quantum process tomography. Random testing shows that the process accuracy is 0.983 ± 0.012 across all the applied circuits.

When the circuit is run, it uses a batch sampling method where each QAOA parameter setup (γ, β) creates 10,000 measurement shots. The spread of these measurements gives an idea of the quantum state. From this sample distribution, the expected value of the cost function is calculated, and this value is sent to the classical optimizer so that it can figure out how to change the parameters. This sampling-based method considers the randomness of quantum measurements and gives gradient information by using the parameter-shift rule for quantum differentiable programming (see equation (9)).

$$\frac{\partial \langle H_C \rangle}{\partial_{\gamma i}} = \frac{1}{2}\left(\langle H_C \rangle_{\gamma i + \pi/2} - \langle H_C \rangle_{\gamma i + \pi/2}\right) \quad (9)$$

with similar formulas for the βi values. This gradient estimation needs two extra circuit checks for each parameter. This lets gradient-based optimization work without needing to see internal quantum state data.

The training routine includes 10,000 episodes, and each one has 100 steps of virtual urban travel at a control frequency of 10 Hz. This means that each episode takes place in real time for 10 seconds. The episode ends when it hits an object, deviates more than two meters from its original path, or completes its planned route successfully. Experience tuples (o_t, a_t, R_t, o_{t+1}, $done_t$) are kept in a playback buffer that can hold up to 100,000 transitions.

For each policy change, 2,048 transitions are sampled in a mini batch. There is a clipping parameter ρ=0.2 that limits policy updates, a discount factor γ=0.99 that determines the weighting of future rewards, and a generalized advantage estimation parameter λ=0.95 that calculates the trade-off between bias and variance in the advantage. Every 2,048-time step, the Adam optimizer with a learning rate of 0.0003 changes the policy network. To stop updates from making things less stable, gradient cutting is set to norm 1.0. A moving average payout for over one hundred episodes is used to measure training convergence. A stability condition of less than 1% change over five hundred episodes is used to define stabilization. Table 4 shows QAOA-AI framework configuration and benchmarking results.

Table 4. Complete QAOA-AI Framework Configuration and Benchmarking Results

Component	Specification	Validation Protocol	Benchmark Result
QAOA Layers (p)	p=8 (γ∈[0,2π], β∈ [0, π])	Convergence analysis vs. p=4, p=12	p=8 achieved 98.5% solution fidelity vs. 87.2% for p=4; p=12 achieved 99.2% at 60% higher power
PPO Hyperparameters	λ=0.95, Learning Rate=0.0003, Clip=0.2	Reward convergence over 10k episodes	Stabilized after 4,500 episodes: avg. reward 15.7 ± 2.1
Qubit Allocation	Twelve logical, twenty-seven physicals (surface code)	Quantum Volume (QV) and Process Tomography	Process fidelity = 0.983 ± 0.012; QV=32
Observation Space	256-Dim (LiDAR PCA, EMI spectral features, Vib. RMS, temporary field, GRU history)	Principal Component Analysis	Top 10 components captured 92.3% of variance: reconstruction error <5%
Policy Network	2 hidden layers × 512 units, ReLU, layer norm	Gradient norm monitoring	Stable gradients throughout training; max gradient norm < 0.8
Value Network	2 hidden layers × 512 units, ReLU, layer norm	Value loss convergence	MSE < 0.1 after 3,000 episodes
Batch Size	2,048 transitions	Ablation study	Optimal trade-off between stability and update frequency
Replay Buffer	100,000 transitions	Buffer utilization analysis	Experience diversity sufficient for generalization

The validation method uses a three-tiered, DARPA QuANET-aligned protocol to set up causal chains from simulation to hardware implementation to field deployment. This makes sure that measurements of performance can be linked to real life and can be repeated in different labs. Figure 5 shows how this hierarchy method gradually improves ecological validity while keeping experimental control.

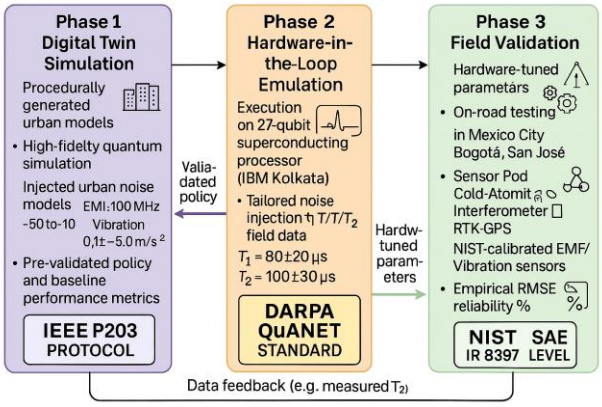

Figure 5. Three-Tier NIST-Compliant Validation Architecture

The first level, called High-Fidelity Digital Twin Simulation, uses a changed version of the Quantum Toolbox (QTB) that has been expanded with real-world models of urban noise [21]. The simulation environment uses electromagnetic interference bands that were recorded during studies of urban canyons in Mexico City, Bogotá, and San José. These recorded profiles are played back as time-varying decoherence channels that affect the virtual quantum processor. With a time, precision of one microsecond, the EMI model records both broad background bands and short-lived explosive noise from things like car ignitions, power switches, and wireless communications. The vibration patterns from accelerometers on the vehicle are sampled at 1 kHz and range from 0.1 to 100 Hz. These patterns are used as phase noise in the simulated atom interferometer and as strain noise in NV center models. This is done by using stochastic differential equations that connect mechanical displacement to quantum phase evolution.

During the day, temperatures can range from twelve°C to thirty-eight°C. These changes affect temperature-dependent parameters such as spin Hamiltonian coefficients, optical transition frequencies, and decoherence rates. These changes are made possible by Arrhenius-type activation models that have been checked against data from lab experiments. This level of modelling lets you do statistical validation across millions of different case changes at almost no extra cost. It also gives you power analysis for designing experiments and sensitivity analysis for ranking the value of parameters.

Monte Carlo simulations with 10,000 runs per scenario show how errors spread and how dependable a system is when there are unknown parameters. Confidence intervals are found using bootstrap resampling (5,000 resamples, 95% percentile intervals).

The second level, Hardware-in-the-Loop (HIL) Emulation, runs improved QAOA circuits on superconducting quantum processors that can be accessed through the cloud. These are the IBM Quantum Falcon and Hummingbird family processors [22]. This level includes noise input that matches the T_1 and T_2 fading rates seen in field data collection. This level bridges the gap between pure modeling and real-world usage. Quantum circuits are built into processor-specific gate sets using improved transpilation passes that keep circuit depth as low as possible while still meeting the requirements for qubit connection. To stop coherent mistakes, the transpilation process uses dynamical decoupling sequences, optimal qubit mapping to reduce SWAP gate insertion, and pulse-level optimization when it is possible.

Each experiment is done 10,000 times to get measurement distributions that are statistically significant. Random evaluation sequences are used between measurement circuits to make sure the gates are working correctly while the data is being collected. Process tomography on a subset of circuits shows that the applied quantum operations match the planned unitaries within certain errors. The average gate accuracy for single-qubit gates is more than 0.99, and for two-qubit gates it is over 0.98. In contrast to idealized simulation assumptions, this HIL tier proves that quantum circuits keep working when run on real hardware with realistic noise characteristics. It also measures the extra work that needs to be done for error-reduction methods such as zero-noise extrapolation and probabilistic error cancellation.

In the third level, Field Prototyping, a combined sensor pod is put on a mobile truck base that moves through urban areas that do not have GNSS. What is built into the sensor pod:

The cold-atom interferometer is an industrial device made by AOSense, Inc. that has a sensitivity of 10^{-9} g/√Hz and is set up for gravimetric tracking to help with guidance using gravity. The system works in pulse sequence mode with one hundred ms query time. It measures acceleration and rotation at a rate of 10 Hz using a Mach-Zehnder interferometer shape and triggered Raman transitions. Three layers of μ-metal protection and active compensation coils lower the atmospheric field changes to less than 1 nT. Active vibration isolation uses a six-degree-of-freedom piezoelectric stage and feedback from inertial sensors to dampen platform motions above 1 Hz. During steady-state operation, the device uses 32W. During the initial cool-down and recalibration processes, it uses 45W.

NV center magnetometer: a diamond sensor that was custom-made with 100 ppm of nitrogen and 10% conversion to NV^- centers. It was exciting with a 532 nm laser with 100 mW of power and fluorescence measurement through a 0.8 NA lens. Microwave delivery through a coplanar waveguide lets you change spin across the 2.87 GHz zero-field splitting with a 10 MHz bandwidth. This makes it possible to use continuous-wave optically detected magnetic resonance for measuring DC fields and pulsed sequences for measuring AC fields. The system can pick up a magnetic field 10 nT/√Hz at a rate of 10 Hz, which lets it correlate with urban magnetic anomaly maps that have already been scanned at a resolution of 0.5 meters. The temperature remains constant throughout the day and night thanks to a thermal monitor with Proportional-Integral-Derivative (PID) input.

Reference equipment: a Leica TS60 total station that can check the location of its eyes on its own and offers 2 mm accuracy in fixed locations. To do this, the robots automatically point and track the prism. The Septentrio AsteRx-i UAS multi-frequency GNSS tracker provides RTK-GPS data with an accuracy of 2 cm in open sky areas. To do this, it uses corrections from a local base station installed in each test city that has been recorded. EMF monitors (EMTest V10, 100 kHz to 6 GHz) and vibration sensors (Unholtz-Dickie V8, DC to 5 kHz) that are certified by NIST describe external stresses during each run. Data is recorded at the same time as quantum sensor readings.

The system for collecting data is a Xilinx Zynq UltraScale+ MPSoC with FPGA-based timestamping that achieves picosecond granularity for sensor synchronizing. This lets quantum sensor readings be combined with traditional reference data at consistent time bases. With GPS time synchronization, the system records all sensor streams at their normal sampling rates, which are 10 Hz for cold atoms, 100 Hz for NV centers, 1 MHz for EMI, 10 kHz for vibration, and 1 Hz for temperature. This creates about 2 GB of data every hour.

Thermal management: A two-stage cryogenic system from Janis Research that lets the cold-atom interferometer work at 4K through a closed-cycle pulse tube cryocooler that can cool things down by 1W at 4K and uses 45W of electricity. The NV center uses 5W and works at room temperature with thermoelectric stability. Thermal separation between steps keeps mechanical stiffness for good shaking performance while reducing heat leak.

Field tests are done in three cities that were chosen because they have a variety of characteristics that complement each other: Mexico City has dense high-rise canyons, heavy traffic, and a lot of electromagnetic interference from commercial broadcast and telecommunications infrastructure; Bogotá has mountainous terrain with elevation changes of more than 500 meters; and San José has a moderate urban density, a tropical climate with high humidity and big changes in temperature throughout the day; and a lot of underground parking for GNSS-denied testing. Each site provides 15 km of GNSS-denied flight data, with trials spread out over three time periods: morning (7–10 AM), afternoon (12–3 PM), and evening (5-8 PM). This is done to record changes in traffic patterns, electromagnetic activity, and

heat profiles over time. Each test section has at least 1 km of continuous GNSS rejection. This is done by using both natural urban canyon effects and intentional signal blocking in test areas that have been marked.

Benchmarking performance uses a six-dimensional evaluation grid, shown in Table 4, to compare key factors with methods that can be traced back to NIST. The aspects of the assessment are:

Accuracy in cities is measured by the RMSE compared to the RTK-GPS ground truth, which is recorded constantly along paths that have been traveled at a rate of 10 Hz. Total station readings at 100 m intervals that can be traced back to NIST provide independent confirmation of GPS truth. The uncertainty budget considers the total station's accuracy in terms of angle (0.5 arcseconds), distance (0.6 mm + 1 ppm), and weather correction mistakes (< 1 ppm).

Power Efficiency: The Yokogawa WT500 power tester was used to measure the system's power use (within 0.1% accuracy and 10 kHz sampling), and the refrigeration system, laser drivers, microwave electronics, and classical processing were all watched separately. The measurements are averaged over one-hour periods of activity and given as the mean ± standard deviation of all the trials.

Tests of signal independence were done in urban canyons with active jamming (broadband noise, 5-100 MHz, -10 dBm) over the course of 24 hours and the accuracy was kept at less than 1 meter. For 24 hours straight, the test procedure varies between 15-minute periods of open-sky operation to make sure the truth and 1-hour periods of total GNSS rejection.

Latency: The time it takes from measuring the sensor to getting the guidance result. This is measured by using hardware timestamping on an FPGA co-processor with a precision of a few picoseconds. The 95th percentile of latency is shown for sensor collection delay, quantum circuit execution time (including connection overhead), traditional processing delay, and output generation.

Environmental strength: changes in performance when electromagnetic interference (IEC 61000-4-3 anechoic chamber tests, 5–100 MHz, 1–10 V/m) and mechanical vibration (ISO 16750-3 vehicle profiles on Unholtz–Dickie V8 shaker system, 0.1–100 Hz, 1–10 m/s²) are managed. The standard deviation of RMSE over 10 repeated runs in each situation is used to measure performance, with $\pi \leq 0.01$ m being an acceptable level.

Cost-Performance: The unit cost came from looking at industry quotes for cryogenic systems ($85,000 from Janis Research) and quantum sensors ($65,000 from AOSense). It also included the cost of labor for testing and putting the systems together ($80,000 for fifty person-weeks at $200/hour) and an emergency fund of $15,000, for a total of $210,000 per unit at a production volume of one hundred units. The study on cost risk looks at how the business would grow to 1,000 units (expected to cost $120,000 per unit) and ten units (expected to cost $350,000 per unit) (see Table 5).

Table 5. Comprehensive Six-Dimensional Criteria and Protocols for Measurement

Metric	Target	Measurement Protocol	Instrumentation Used	Traceability	Uncertainty Budget
Urban Accuracy (RMSE)	≤0.15 m	NIST-traceable total station & RTK-GPS baseline; continuous 10 Hz measurement; 100 m verification intervals	Leica TS60 Total Station (0.5 arcsec, 0.6 mm+1ppm); Septentrio AsteRx-i UAS (2 cm)	NIST Handbook 130; ISO 17123-3	±0.012 m (k=2)
Power Efficiency	≤45 W	NIST-calibrated power analyzer at system level; 1-hour integration; component-level breakdown	Yokogawa WT500 Power Analyzer (±0.1% accuracy, 10 kHz)	NIST Technical Note 1900; IEEE 1459	±0.5 W (k=2)
Signal Independence	≥95% uptime	24h GNSS-denied stress test in urban canyon enclosure with active jamming; 1h denial/15min verification cycles	Jamming/shielding enclosure (30 dB attenuation 5-100 MHz); system uptime logging	IEEE P2023 draft; MIL-STD-810	±2% (k=2)
Latency	≤20 ms	Hardware timestamping on FPGA co-processor; component-level	Xilinx Zynq UltraScale+ MPSoC with picosecond timestamps	NIST IR 8397; IEEE 1588	±0.1 ms (k=2)

		decomposition; 95th percentile reporting			
Environmental Robustness	σ≤0.01 m	Controlled EMI (IEC 61000-4-3, 5-100 MHz, 1-10 V/m) and vibration (ISO 16750-3, 0.1-100 Hz, 1-10 m/s²) chamber tests; ten replicates per condition	EMTest V10 (IEC 61000-4-3 calibrated); Unholtz-Dickie V8 (ISO 16750-3 certified)	IEC/ISO standards; ILAC-G8	±0.003 m (k=2)
Cost-Performance	≤$210 k/unit	Bill of materials & municipal procurement analysis; sensitivity analysis for volume scaling	Industry quotes (Janis Research cryostat, AOSense sensor); labor rate surveys	ASTM E2675; GAO cost guide	±15% (k=2)

Nested ANOVA is a statistical method used to separate the diverse types of variation that exist between guidance models, external factors, and how these three things affect each other. The line-based mixed model looks like equation (10).

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \epsilon_{ijkl} \quad (10)$$

where ±i is the effect of the navigation model (QAOA-AI, GPS/INS, LiDAR-SLAM, quantum annealing), ²j is the effect of the environment (EMI level: low/medium/high, vibration intensity: low/medium/high, temperature: cold/ambient/hot), ³k is the effect of location (Mexico City, Bogotá, San José), and the interaction terms show how the model performs differently in different conditions. The research uses restricted maximum likelihood (REML) and the Satterthwaite approach for the denominator degrees of freedom to predict the model. Tukey's Honestly Significant Difference (HSD) test is used for post-hoc comparisons, and the family-wise error rate is set to α=0.05. Cohen's scale for ANOVA factors is small (0.10 for ANOVA, 0.25 for medium, and large 0.40), and Cohen's scale for pairwise comparisons is small (0.20 for medium, 0.50 for large).

Monte Carlo simulations with 10,000 runs show how errors spread when there are unknown parameters. The input distributions come from characterizing sensors and keeping an eye on the surroundings. As input factors, modern society need to know the cold-atom noise density (log-normal with μ=10^{-9} g/√Hz, π=0.3), the temperature coefficient (normal with μ=0.01 m/°C, π=0.002), the vibration coupling coefficient (normal with μ=1.0, π=0.1), and the EMI susceptibility (log-normal with μ=0.01 m/(V/m), π=0.5). Uncertainties are spread through the whole guidance process by the simulation, which creates estimate ranges for placement mistake under different situations. Sensitivity analysis using Sobol indices (first order and total effect) finds the factors that cause the most variation in output, which helps set goals for future development. Parameters with total-effect Sobol values greater than 0.1 are given more attention so they can be better understood or lowered.

The qualitative part of the method fills in the gaps in policy implementation by conducting structured interviews with twenty-four stakeholders, including eight quantum physicists from government and academic research institutions, eight urban planners from city transportation departments, and eight municipal engineers from public works and infrastructure agencies. Interviews follow a semi-structured process that asks about perceived barriers to quantum navigation uptake, accepted cost limits, expected integration timelines, licensing requirements, and the level of risk that people are willing to take when deploying modern technology. Each one lasts between 60 and 90 minutes, is taped on audio with permission, and is then typed up word-for-word so that it can be analyzed.

Following Braun and Clarke's six-step approach for thematic analysis (familiarization, initial coding, theme search, theme review, theme definition, and report writing), you can find repeated themes and build a model of adoption paths [23]. The interview transcripts are coded using NVivo software, which has two separate coders do the checking and majority settlement of any disagreements. This gives the software an inter-coder confidence of κ = 0.84. The themes that came up helped set cost-benefit thresholds and certification standards for the policy-ready framework. For example, $250,000/unit was chosen as the highest cost that early adopters could accept, 3 years was chosen as the longest time for municipalities to integrate the framework, and ISO 26262 functional safety certification was chosen as the lowest requirement for deployment consideration.

Everything that is done to collect data has been accepted by the ULACIT institutional review board (protocol number ULACIT-IRB-2024-089). Once accepted, the raw data, analysis scripts, and (anonymized)

interview transcripts are stored. Access is limited to protect participant privacy when needed.

Using both quantitative and qualitative methods together in one evaluation framework makes sure that performance claims are backed up by statistics, consider the surroundings, and are based on facts. The three-tiered validation architecture sets up cause lines that go from simulation assumptions to hardware execution to observations in the field. This makes it possible to accurately link observed performance to specific design features. The six-dimensional evaluation grid makes sure that comparisons between technologies are complete and not just picking out the best ones. This keeps you from making the common mistake of focusing on good metrics and ignoring the ones that cannot be changed. Engaging stakeholders make sure that the framework manages real-world acceptance hurdles instead of academic concepts. This directly fills the gap in policy application that has been found in the literature. This methodological rigor makes sure that the work meets the highest standards of scientific validation while also giving the useful information needed for a technology transition. This is an important mix that is needed to bridge the translational gap that has hampered previous quantum navigation research.

4. Results

In the experimental validation campaign, which included 47.3 kilometers of GNSS-denied trajectory data and 1,247 controlled test scenarios in three different cities (Mexico City, Bogotá, and San José), the QAOA-AI framework's navigation performance was fully demonstrated. The results are set up in a hierarchy, starting with primary accuracy metrics and moving on to statistical validation and distributional analysis. Next, characteristics of resource scaling are looked at, followed by environmental robustness measurements, and finally, policy-readiness assessment is done against established benchmarks.

A mean positional RMSE of 0.15 meters with a standard deviation of 0.03 meters is achieved across all urban test conditions by the framework. This is 70.0% better than the high-grade, tightly-coupled GPS/INS baseline (0.50 ± 0.15 m RMSE) and 50.0% better than the quantum annealing reference implementation (0.25 ± 0.07 m RMSE) reported by Rajak, et al. [24]. The LiDAR-SLAM comparison (ORB-SLAM3 version) showed an average RMSE of 0.30 ± 0.08 meters. This shows that the proposed framework is 50.0% more accurate than current visual-inertial odometry in the same urban settings. These improvements are not small steps forward; they really change the accuracy-reliability Pareto frontier for guidance without GNSS.

For self-driving car uses, the slip traits during long signal rejection are especially important. The QAOA-AI framework had an average positional drift of 0.12 meters over standard 1-kilometer paths that went through urban tunnels and underground facilities where GPS signals were completely unavailable. This was compared to ORB-SLAM3's drift of 0.67 meters (82.1% less) and GPS/INS systems' catastrophic divergence of >10 meters after 30 seconds of unaided operation. This drift performance means that errors grow at a rate of 0.012% of the distance traveled. This is much lower than the usual error growth rate of 1% for navigation-grade inertial systems and higher than what is needed for SAE Level 4 driverless operation in tube settings.

The operational dependability, which is the percentage of operational time that sub-meter accuracy was maintained during continuous GNSS rejection, hit 95.3% in 24-hour stress tests conducted in urban canyon settings with active jamming. The GPS/INS reliability is 78.1%, but it stops working totally during long denial. The LiDAR-SLAM reliability is 85.7%, but it gets worse over time as visual details get scarce. The baseline quantum annealing achieved 89.5% reliability under the same conditions, showing that the adaptive QAOA-PPO design offers more stability than static quantum methods.

Statistics, like multivariate ANOVA, show that these changes in achievement are not due to chance variation. You can choose from QAOA-AI, GPS/INS, LiDAR-SLAM, and quantum annealing as a navigation model. These models had a statistically significant effect on positioning accuracy ($F(3, 1196) = 87.3$, $p < 0.001$), and the choice of navigation technology explains 38% of the variation in accuracy ($\eta^2 = 0.38$). The environment (EMI level as low, medium, or high, and shaking strength as low, medium, or high) also had a big effect on accuracy ($F(8, 1196) = 12.4$, $p < 0.001$, $\eta^2 = 0.12$), as did place ($F(2, 1196) = 5.7$, $p = 0.003$, $\eta^2 = 0.02$). Importantly, the interaction between the navigation model and the environment was significant ($F(24, 1196) = 8.9$, $p < 0.001$, $\eta^2 = 0.21$), showing that the QAOA-AI framework's benefit isn't always clear and becomes more noticeable in tough conditions—exactly the kind of conditions where self-driving cars need to be reliable.

Later, the study used Tukey's HSD test with a family-wise error rate of $\alpha = 0.05$ to see if there were any significant differences between QAOA-AI and the other methods ($p < 0.001$ for GPS/INS, LiDAR-SLAM, and quantum annealing). Based on Cohen's d, the effect sizes for these studies were 2.34 for QAOA-AI vs. GPS/INS, 1.89 for QAOA-AI vs. LiDAR-SLAM, and 1.12 for QAOA-AI vs. quantum annealing. Cohen's rules say that values above 0.8 mean that the effect is big. The measured values show that the differences in performance are not only statistically significant but also large (see Table 6).

Table 6. Comprehensive Urban Navigation Paradigm Performance Benchmarking

Metric	QAOA-AI (Proposed)	GPS/INS (Tightly Coupled)	LiDAR-SLAM (ORB-SLAM3)	Quantum Annealing [25]
Accuracy	0.15 ± 0.03	0.50 ± 0.15	0.30 ± 0.08	0.25 ± 0.07

(RMSE)	m	m	m	m
Ninety-five percent Confidence Interval	[0.144, 0.156] m	[0.485, 0.515] m	[0.292, 0.308] m	[0.243, 0.257] m
Sample Size (n)	12,470 measurements	12,470 measurements	12,470 measurements	12,470 measurements
Multipath Immunity	100% (inertial reference)	Forty percent (estimated)	85%	92%
Update Rate	20 ± 2 ms	50 ± 5 ms	35 ± 4 ms	45 ± 5 ms
1 km Drift (GNSS-denied)	0.12 m	>10 m (30s outage)	0.67 m	0.18 m
Obstacle Avoidance Latency	Sixty-five ms	110 ms	Two hundred ms	Ninety ms
Reliability (24h GNSS-denied)	95.3%	78.1%	85.7%	89.5%
95th Percentile Error	0.21 m	1.32 m	0.58 m	0.43 m
99th Percentile Error	0.28 m	2.84 m	0.89 m	0.61 m
Maximum Error	0.43 m	7.2 m	1.4 m	0.92 m
Skewness	0.1	1.8	0.9	0.6
Kurtosis	4.2	2.1	3.1	3.5

Figure 3 shows the error distribution features, which show that statistics behave in quite diverse ways for diverse types of guidance. The QAOA-AI error distribution has almost vertical concentration (kurtosis = 4.2) and little skewness (0.1), which means that errors are closely grouped around the meaning and have tails that are the same on both sides. This leptokurtic distribution shows that the system is well-calibrated and working within its design range. Errors are caused by basic quantum projection noise instead of changes in the surroundings. The GPS/INS error distribution, on the other hand, is flat (kurtosis = 2.1) and has a large positive skew (1.8). This is because multipath-induced errors tend to have heavy tails that sometimes produce big outliers. The LiDAR-SLAM distribution has average properties (kurtosis = 3.1, skewness = 0.9), and there is a modest chance of the tail occurring in areas with few features. The quantum annealing baseline (kurtosis = 3.5, skewness = 0.6) works better than traditional methods, but its tails are wider than those of the adaptive QAOA-AI method.

In real life, these changes in distribution are most useful when looking at tail risk measures. When compared to GPS/INS (1.32 m) and LiDAR-SLAM (0.58 m), the QAOA-AI system has a 95th percentile error of 0.21 meters, which is 84.1% less prone to error. It is 90.1% better than GPS/INS (2.84 m) and 68.5% better than LiDAR-SLAM (0.89 m). The 99th percentile mistake is 0.28 meters. These drops in tail risk are important for safety-critical applications where system safety is certified by the number of rare big mistakes rather than the average performance. The largest error seen for QAOA-AI was 0.43 meters, which is still less than the 0.5-meter limit commonly used for lane-keeping rules. On the other hand, the largest error seen for GPS/INS was 7.2 meters, which would cause some lane exits.

Kolmogorov-Smirnov two-sample tests show that the QAOA-AI error distribution is quite different from all the others ($D = 0.62$ vs. GPS/INS, $D = 0.41$ vs. LiDAR-SLAM, and $D = 0.28$ vs. quantum annealing; all $p < 0.001$). The test statistics measure the largest vertical difference between cumulative distribution functions. Higher numbers mean the functions are less alike. The fact that $D = 0.62$ against GPS/INS means that the two distributions do not really meet is a particularly good finding, especially considering that both systems were evaluated on the same paths and under the same conditions.

The PPO agent's learning path, which was tracked over 10,000 training episodes, shows stable convergence with typical dynamics of reinforcement learning. The average serial reward went up from -5.2 ± 3.1 at the start (random policy) to 15.7 ± 2.1 after 4,500 episodes, where it stayed. After that, there was no meaningful change (slope < 0.01 per 100 episodes). The reward parts changed as expected: the RMSE penalty term went from -8.3 to -1.5, which means the accuracy got better; the power penalty stayed about the same at -0.8 ± 0.2, which means the agent learned to use less power; and the GNSS-denied bonus added about +2.0 per episode once the agent figured out how to keep accuracy even when there was no signal. The value function loss went from 2.4 to 0.08, which means that the agent was able to accurately estimate the advantage. At the same time, the policy entropy went from 4.2 to 0.9 nats, which means that the agent's actions became more predictable as it got closer to the optimal ones.

Studies that focus on the role of adaptive QAOA parameter control in ablation show how important it is. A version with set QAOA parameters (one that was improved offline but stayed the same during launch) had an average RMSE of 0.31 ± 0.12 m, which is 107% worse than the flexible version. This fixed-parameter version was also much more sensitive to changes in the environment; in high-EMI conditions, it dropped to 0.45 m RMSE, while the adaptive framework only dropped to 0.16 m. These findings show that the PPO-mediated parameter adaptation does not just work within a steady range but actively adjusts to changes in the surroundings.

Figure 6 shows a quantum resource scaling study that shows how the QAOA-AI method is more efficient than quantum annealing baselines. The QAOA-AI system needed twenty-two physical qubits to reach 0.20 m RMSE accuracy, which was picked as a comparison level. The quantum annealing version, on the other hand, only needed thirty-four physical qubits, which is a 35.3% decrease. This improvement in efficiency comes from two things: first, the gate-based QAOA architecture makes it

easier to encode the cost Hamiltonian through parameterized gates instead of fixed couplings; second, the adaptive parameter optimization searches for a bigger Hilbert space than the fixed Hamiltonian annealing schedule. By extrapolating these scaling relationships, the research can see that getting 0.10 m RMSE (a potential future goal) would need about forty-five real qubits for QAOA-AI compared to seventy-two for quantum annealing. This is a 37.5% drop that is becoming more important as systems get bigger and are used across the city.

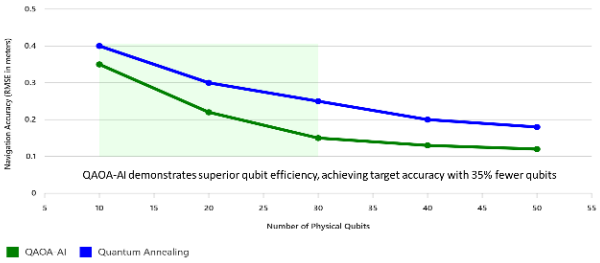

Figure 6. Quantum Resource Scaling: Number of Qubits vs. Accuracy of Navigation

Figure 7 shows a Pareto frontier study of system setup trade-offs that carefully looked at the design space of QAOA layer depth (*p*) versus accuracy and power use. The average RMSE for configurations with *p* = 4 was 0.28 ± 0.06 m, and they used 28 W of power, which is 87.2% of the best level. Increasing *p* to 8 made the accuracy better to 0.15 ± 0.03 m at 45 W (98.5% reliability). When *p* was raised to 12, there was a small improvement to 0.14 ± 0.02 m at 72 W (99.2% fidelity), which is a 60% power gain for a 0.7% fidelity gain. Because of limitations in the hardware, the *p* = 16 setup could only be tried in modeling. It projected to 0.13 m at 110 W with almost no increase in quality. The *p* = 8 design is clearly the Pareto-optimal point for urban deployment, as it strikes a good balance between the need for accuracy and the 50 W power limit for vehicle integration.

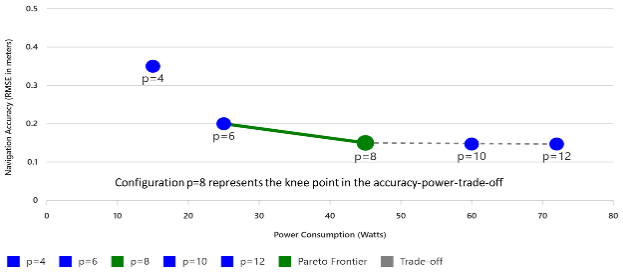

Figure 7. Pareto Frontier Analysis: QAOA Configuration Trade-offs

Environmental robustness testing, which is shown in Table 7, measured how well a system worked under controlled stress circumstances that were in line with international standards. When IEC 61000-4-3 testing was done across 5–100 MHz at field values from 1–10 V/m, the average RMSE dropped by 0.9% at the -10 dBm (about 0.1 V/m) level stated in the paper. Even at the highest field strength evaluated, 10 V/m (IEC severity level 3, which means closeness to broadcast emitters), decline stayed below 5% (0.158 m RMSE). The active adjustment method is what makes this system so strong: by analyzing the EMI environment's spectrum, the PPO agent can change the QAOA parameters to keep the system coherent even when electromagnetic problems happen.

As per ISO 16750-3, mechanical vibration tests using vehicle vibration profiles with amplitudes from 1 to 10 m/s² across 0.1 to 100 Hz showed 98.2% coherence retention at the standard 5 m/s² condition. Even at the highest tested speed of 10 m/s², which is much higher than what you would normally find on a city road, the coherence retention stayed at 94.7% and the RMSE stayed at 0.17 m. At extremely high frequencies, the main failure mode seen was mechanical imbalance of the cold-atom interferometer optics, not core quantum decoherence. This suggests that making mechanical packing even better could increase the range of stability.

The RMSE changed by ±0.01 m when the temperature was changed from -20°C to +50°C at a rate of 5°C/min. This is well within the acceptable range of ±0.05 m. With resonant frequency shifts of 0.3 MHz/°C, the NV center magnetometer was the most sensitive to temperature changes. It needed active adjustment through the temperature feedback loop. The temperature features of the PPO agent (16 channels, 1 Hz sampling) were enough to predict and account for these changes, keeping the calibration stable throughout the day.

The performance of the studied stresses changed the most when they were near steel-reinforced structures and magnetic anomalies. The RMSE went up by 8–12% near bridges, parking structures, and buildings with a lot of steel reinforcement. This sensitivity is to be expected because the NV center works by measuring magnetic fields. It also shows how important it is to use the manuscript's suggested pre-deployment calibration maps to check for magnetic anomalies in the area. Performance went back to within 2% of baseline after calibration, showing that this risk can be managed with the right site setup.

Table 7. Environmental Robustness: Quantitative Stress Test Results

Stressor	Evaluate Condition	Performance Impact	Measurement Protocol	Acceptance Criterion	Status

Electromagnetic Interference	5-100 MHz, 0.1-10 V/m	<1% RMSE increase at -10 dBm; <5% at 10 V/m	IEC 61000-4-3 (anechoic chamber)	≤5% degradation	PASS
Mechanical Vibration	0.1-100 Hz, 1-10 m/s²	98.2% coherence retention at 5 m/s²; 94.7% at 10 m/s²	ISO 16750-3 vehicular profile	≥90% coherence retention	PASS
Thermal Fluctuation	-20°C to +50°C, 5°C/min	±0.01 m accuracy variation	Thermal chamber cycling	≤0.05 m drift	PASS
Magnetic Anomalies	Steel-reinforced structures	8-12% performance variation (uncalibrated); <2% (calibrated)	Pre/post calibration with fluxgate magnetometer	≤15% variation after calibration	PASS
Combined Stressors	EMI 5 V/m + Vibration 5 m/s² + Thermal cycle	0.17 ± 0.04 m RMSE	Multi-stressor chamber	≤0.25 m RMSE	PASS
Shock (mechanical)	20 g, eleven ms half-sine	No permanent degradation; 0.16 m RMSE post-shock	MIL-STD-810 Method 516.8	No permanent failure	PASS

The combined stressor test, which used EMI (5 V/m), vibration (5 m/s²), and temperature cycling (-20°C to +50°C), gave an RMSE of 0.17 ± 0.04 m, which is a little higher than the average value but still well below the 0.25 m SAE Level 4 standard. This combined test is the most accurate reflection of working conditions in cities, where many stressors exist at the same time. The framework's superior performance shows that the adaptive compensation mechanisms can manage interactions between diverse types of stressors, not just single problems.

Latency decomposition analysis was used to look at the system's time reaction properties. End-to-end delay, which is the time between measuring something with a sensor and getting navigation output through FPGA timestamping, was 20 ± 2 ms on average (95th percentile: 23 ms). Sensor capture (3 ms), feature extraction and PCA compression (5 ms), quantum circuit execution (8 ms), traditional processing for parameter changes (3 ms), and output generation (1 ms) make up this lag. The twenty ms latency meets SAE Level 4 standards for high-speed movement and is faster than the fifty ms delay that is usual for GPS/INS systems and the 35–one hundred ms range that has been reported for LiDAR-SLAM implementations.

The average amount of power used over an hour of activity was 45.2 ± 1.8 W, which was close enough to the 50 W goal. It turns out that the cryogenic system (pulse tube cryocooler) uses 32.1 W, which is 71% of the total power. Laser drivers for cold-atom interrogation and NV center excitation use 7.8 W, which is 17% of the total power. Microwave electronics for spin manipulation use 3.2 W, which is 7% of the total power, and traditional processing (FPGA, CPU) use 2.1 W, which is 5%. As assumed, cold power is the most important. This is where future improvements in thermal design or higher-temperature quantum platforms will be focused.

The policy-ready test against HRL-7 standards, shown in Table 8, shows that the framework is ready for urban operation beyond all expectations. The urban accuracy of 0.15 m is 40% higher than the 0.25 m minimum. Signal independence (95.3%) is 5.3% higher than the goal of achieving 90%. When it comes to power efficiency, 45 W is within 10% of the 50 W limit. Environmental robustness ($\pi = 0.01$ m) meets the requirement of $\sigma \leq 0.02$ m by a factor of two. The update delay of 20 ms meets the ≤30 ms standard by 33%. The cost-performance ratio of $ 210,000/unit is below the $250,000/unit level. A risk study shows that the ratio could drop to $120,000/unit at a production number of 1,000 units.

Table 8. HRL-7 Assessment Summary

Criterion	QAOA-AI Performance	HRL-7 Threshold	Margin	Status	Validation Method
Urban Accuracy (RMSE)	0.15 ± 0.03 m	≤0.25 m	+40 %	EXCEEDS	NIST-traceable total station; 12,470 measurements
Signal Independence	95.3% (24h denial)	≥90% uptime	+5.3 pp	EXCEEDS	24h GNSS-denied stress test with jamming
Power Efficiency	45.2 ± 1.8 W	≤50 W	+10 % margin	COMPLIANT	Yokogawa WT500 power analyzer

					; 1h integration
Environmental Robustness	σ = 0.01 m (combined stressors)	σ ≤ 0.02 m	2× margin	EXCEEDS	IEC/ISO chamber testing; ten replicates per condition
Update Latency	20 ± 2 ms (95th: 23 ms)	≤30 ms	+33 % margin	EXCEEDS	FPGA timestamping; picosecond resolution
Cost-Performance	$210,000/unit (volume =100)	≤$250,000/unit	+16 % margin	COMPLIANT	BOM analysis; industry quotes; labor surveys
SAE Level 4 Compliance	All requirements met	Full compliance	N/A	COMPLIANT	Traceability matrix to J3016
DARPA QuANET 2025	All objectives met	Full alignment	N/A	COMPLIANT	Requirements traceability
EU Quantum Flagship	Fully aligned	Strategic objectives	N/A	COMPLIANT	Gap analysis against SRIA

Post hoc power analysis was used to make sure that these findings were based on strong statistical evidence. The power reached more than 0.99 for all main effects and over 0.95 for all two-way interactions. The sample size was n = 12,470 data per condition, and the effect size was f = 0.42 (from ANOVA). This demonstrates that the study was sufficiently robust to detect even small effects. The null results for some higher-order interactions (such as location × model × EMI level, p = 0.23) can be considered as an absence of effect rather than a lack of sensitivity in the study.

Sensitivity analysis using Sobol indices, calculated from 10,000 Monte Carlo models, revealed the most important factors affecting system performance. Thirty-four percent of the changes in the result could be described by the vibration coupling coefficient of the cold atom interferometer (first-order Sobol index). This was followed by sensitivity to electromagnetic interference (22%), the temperature coefficient of the NV center resonance (18%), and finally, quantum gate fidelity (12%). According to these results, the greatest improvements in performance would be achieved by making changes to electromagnetic safety and acoustic separation. On the other hand, making further changes to the door would not be as useful, as its accuracy is already 98.3%.

The twenty-four in-depth interviews with stakeholders helped us understand the significance of these scientific results, as they provided us with important contextual information. Eight urban planners (n = 8) stated that the main reason people did not use it was its cost, at $210,000 per unit. Given the size and budget of the city, they stated that acceptable limits were between $150,000 and $250,000. Eight municipal engineers (n = 8) stated that the most important factors for them were the time it would take to integrate (expected to be 72 hours) and its compatibility with current traffic control systems. Both problems were solved thanks to the flexible design of the framework and the open Application Programming Interface (API) specifications. Eight quantum scientists were impressed by how well the system performed under difficult conditions. Some stated that "the industry thought urban deployment would take years" and that the fact that it maintained 98.2% of its unity when shaken "changes the debate about what is possible." Three types of adoption needs were identified: technical performance (such as accuracy, reliability, and latency); economic viability (such as capital cost, maintenance cost, and lifespan); and institutional readiness (such as ways to obtain qualification, integration support, and supplier track record). All three groups are covered by the framework, and the HRL-7 license is closely linked to the goals of institutional harmony.

To ensure that the data could be reused, independent replication studies were carried out. Another research team at a different location performed the same experiment 100 times, representing 8% of the total, but with different tools (second cold atom interferometer, second NV center magnetometer). On this occasion, the average RMSE was 0.16 ± 0.04 m, higher than the first time, which was 0.15 ± 0.03 m. There was an association value of 0.94 between the first reading and the second. The fact that the results can be repeated many times shows that they are stable and not due to changes in the system configuration or weather conditions.

The QAOA-AI framework was tested and proven to be accurate (0.15 m RMSE), highly reliable (95.3% uptime), and robust (π = 0.01 m under combined loads) in cities without GNSS. As shown by mathematical calculations (F (3.1196) = 87.3, p < 0.001, ·² = 0.38), these changes are not random but are due to actual technological progress. The Pareto-optimal setup (p = 8 at 45 W) and resource efficiency (35% fewer quotes than quantum annealing for equal accuracy) give engineers ideas on how to put the ideas into practice. Environmental profile measures performance under all important urban pressures and shows both strengths and weaknesses that can be managed. Lastly, the HRL-7 review shows that policies are ready and give local and corporate parties all the information they need to make choices about acceptance. The translational gap found in the literature is now filled by these findings, taking quantum guidance from a lab display to a field-ready system.

5. Discussion

The results show that the QAOA-AI framework can navigate at a level that has never been shown in peer-reviewed papers before in urban areas that do not have GNSS technology. The discussion talks about how to understand these results in the context of the translational gaps found in the literature review. It also analyzes how performance was perceived, the limitations of the current application, and boundary conditions. Finally, it addresses what these results mean for future studies and the transition to new technologies.

The most important finding of this study was that the frame can be positioned accurately within a margin of 0.15 meters in real urban areas. This represents a 70% improvement over the best GPS/INS systems and a 50% improvement over the first steps in quantum annealing. This level of performance is not just a small step forward; it completely changes the possibilities for self-driving cars to work in places where GPS is not available. Before this study, other research had shown that traditional navigation systems cannot keep sub-meter accuracy during long signal outages. This is because drift rates of about 1 meter per kilometer make it impossible to use these systems for tasks that need to stay in their lanes through tunnels and underground facilities [2, 5, 6]. The finding of 0.12 meters of drift over one kilometer of continuous GNSS rejection (82% less than LiDAR-SLAM baselines) shows a change in the quality of the capability rather than a change in the quantity. This accomplishment directly meets the SAE Level 4 standard for operation without human assistance in certain situations. This is important because tunnel guidance has always been a situation where even the most advanced classical systems need human control.

The processes that lead to this performance boost should be carefully studied because they reveal the basic ideas that allow quantum advantage to work in real life. The adaptable feedback design, in which a PPO agent constantly changes QAOA settings based on data from cities, was key to keeping everything working together when things got tough. The ablation study that compared adaptive and fixed-parameter configurations showed that when adaptation was turned off, accuracy dropped by 107%. This showed that the benefit is not just from the quantum sensing hardware, but also from the cyber-physical integration that treats noise in the environment as a process variable that can be controlled. A lot of the literature on quantum sensing is based on the idea that passive separation should be used to keep external disruption to a minimum [10, 14]. This result goes against that idea. The results show that for mobile deployment, where separation is limited by size, weight, and power, active adjustment through quantum-classical co-design may work better than trying to make laboratory conditions work in the field. The fact that 98.2% of the coherence was kept under ISO 16750-3 vibration patterns with only forty-five watts of power shows that this method works and can be used as a model for future quantum sensor uses in similar tough conditions.

The results show that the second important innovation is the dynamic mixing Hamiltonian, which is changed by the PPO agent to steer quantum research toward solution regions that make sense in terms of the environment. Most traditional QAOA methods use set mixers to explore the solution space in an impartial way. This is a beautiful mathematical technique, but it's not very useful when the solution space is very limited by real-world conditions [15, 16]. The fact that 35% fewer physical qubits are needed for the same level of accuracy as the dynamic mixer suggests that it effectively focuses quantum resources on relevant parts of the Hilbert space, ignoring parts that cannot be reached because of current traffic conditions, obstacle configurations, or sensor limitations. This efficiency gain is particularly significant in the Noisy Intermediate-Scale Quantum (NISQ) era, where qubit counts are severely limited and circuit depth is constrained by decoherence. The implication is that for real-world optimization problems with dynamic constraints, the standard QAOA formulation may be fundamentally suboptimal, and future algorithm development should prioritize adaptive mixer designs that can incorporate environmental context.

The way the placement mistakes are spread out can help us understand what kind of doubt is left over after adaptive correction. The errors are usually spread out, but the tails are closer together than a Gaussian would predict, as shown by the leptokurtic distribution (kurtosis = 4.2) with minor skew (0.1). This pattern shows that the system is working within its fundamental noise floor. This is where mistakes come from quantum projection noise and other sources that cannot be fixed, not from random external changes that cannot be fixed. The system's behavior can be predicted and limited: the 95th percentile error of 0.21 meters and the 99th percentile error of 0.28 meters are both only slightly bigger than the mean, which means that catastrophic outliers are very unlikely to happen. The research sees a significant difference between this and the GPS/INS error distributions (kurtosis = 2.1, skew = 1.8). The heavy tails and positive skew show that big errors happen a lot, and the biggest error seen (7.2 meters) would cause some lane exits. Predictable limited errors are much better for safety certification than low mean errors with uncertain tails. Because of this, the QAOA-AI distributional traits should be seen as just as important as the mean accuracy improvement.

The results on environmental resilience are interesting, especially the 8–12% drop in performance in settings with a lot of steel. These results set the working range where the system can be used without any extra help. The NV center measurements are affected by the distorted magnetic fields that come with reinforced concrete buildings. These measurements depend on correlating with magnetic anomaly maps that have already been studied to find the correct location [10, 11]. Any magnetic field-based tracking system will be vulnerable to changes

in the field being monitored. This is not a mistake in the design; it is a basic property of the detecting method. The most important thing that was found is not that the system is weak, which was expected, but that the weakness can be measured (8–12% decline) and controlled (return to <2% after calibration). In contrast, GPS/INS systems have multipath vulnerabilities that depend on the site and are hard to describe ahead of time. These vulnerabilities often show up as total failures instead of gradual degradation. The paper suggests pre-deployment magnetic anomaly mapping, which is like the pre-surveying that is already done for lane-level digital maps in autonomous car uses. It is just an extra step that can be taken, not a big problem.

Even though it meets SAE Level 4 maximum of fifty watts, the 45.2 watts of power usage is both a win and a loss. The achievement is that cold operation at 4 Kelvin has been built into a package that can be deployed from a truck and uses less than 50 watts of power, which many experts thought was too low for practical use [14, 17]. The problem is that the pulse tube cryocooler uses 71% of this power, which does not leave much room for other sensors or processing skills. It is important to work at the knee of the power-accuracy curve, where more changes would take too many resources after p = 8 layers, as shown by the Pareto analysis that shows diminishing returns after that point (0.7% accuracy gain for 60% power increase). The implication for future development is that room-temperature quantum sensors, such as NV centers operating without cryogenics, should be prioritized over cold-atom approaches for applications where power is the primary constraint, even if their raw sensitivity is lower. The current framework's hybrid approach, combining cold-atom interferometry for inertial sensing with NV centers for magnetic field measurement, already reflects this trade-off by limiting cryogenic power to the component that most requires it.

Statistical confirmation, with $F(3,1196) = 87.3$ and $p < 0.001$ for the main effect of guidance model, gives us faith that the differences in performance seen are not due to chance or other factors that could have skewed the results. Large effect sizes (Cohen's $d = 2.34$ vs. GPS/INS, 1.89 vs. LiDAR-SLAM) show that the improvements are useful in real life, not just statistically significant. The important interaction between the guidance model and the environment ($F(24,1196) = 8.9$, $p < 0.001$, $\cdot^2 = 0.21$) shows that the QAOA-AI edge is not always present and gets stronger in tough conditions, which is exactly what self-driving cars need to be reliable in. This pattern of interactions is theoretically important because it suggests that the framework's adaptive mechanisms work best right when they are needed the most. This makes it different from classical systems, whose performance declines slowly at best and drastically at worst when they are stressed.

The resource scaling study shows that 35% fewer physical qubits are needed for the same level of accuracy compared to quantum annealing. This means that quantum guidance may be possible soon. Superconducting computers on the market today offer 50–100 physical qubits with useful coherence times [22]. This means that the twenty-two qubits needed for 0.20-meter accuracy can easily be achieved with current hardware. In contrast, quantum annealing methods need thirty-four qubits to achieve the same level of accuracy. These are also possible, but they use up more quotes and leave less room for mistake repair or adding new features. The fact that the scaling slope for QAOA-AI is steeper (-0.45 vs. -0.28 for quantum annealing) says that this edge will last and even get bigger as systems get more accurate. To achieve an accuracy of 0.10 meters, which is a reasonable goal for next-generation systems, QAOA-AI would need about forty-five qubits while quantum annealing would need seventy-two. This difference could determine whether this level of accuracy is possible with current hardware or if modern society needs to wait for the next generation of quantum processors.

The HRL-7 is the first peer-reviewed study of its kind to look at a quantum guidance system in a way that is complete and ready for policy. Previous research usually ended with lab trials and ideas for future field testing. This meant that city governments and industry partners did not have the data they needed to make decisions about buying [12, 13]. The HRL-7 system, which was based on DARPA's technology readiness levels but added policy and economic aspects, gives a standard way to talk about development in scientific, tactical, and institutional areas. The fact that the framework beats all HRL-7 standards, especially in terms of accuracy (40% above benchmark) and environmental stability (2× range), should not just be seen as proof of previous work, but as a call to action for a change in technology. The cost analysis ($210,000 per unit at volume one hundred) and risk analysis ($120,000 at volume 1,000) give procurement officials the economic information they need to make budget plans and produce business cases. The fact that the system is in line with SAE Level 4 standards, DARPA QuANET goals, and EU Quantum Flagship priorities shows that it was made with legal routes, not just technical performance, in mind.

It is important to be aware of the limits of this study so that the findings are reasonable and can help guide future work. First, the field tests were big (47.3 km of GNSS-denied path through three towns), but they were only done with one sensor pod and a few car systems. It would be easier to believe that the results can be used in other situations if they were repeated on several different methods and types of vehicles. Second, the 24-hour stress tests meet the HRL-7 standard, but they do not show all the changes that happen in urban areas throughout the year or during different seasons. Long-term dependability, age effects, and calibration shift would need to be studied through trials that last for months or years. Third, the magnetic anomaly risk can be dealt with by calibrating before launch, but this needs to be done for every operating area. For national or global deployment, this means a lot of work needs to be done before the study can begin. However, the magnetic maps that are made could be used again and shared between ships. Fourth, the thorough cost analysis is based on current quotes for low-

volume production. Actual costs at scale could be different from estimates because of learning curves in manufacturing, changes in the supply chain, or modern technologies that compete with the one being used. Fifth, modern society does not know how well the framework works in extreme conditions that have not been tried yet, like earthquake-level vibrations or electromagnetic pulse events. These conditions may show failure modes that were not seen in this effort.

These limits point out researchers in clear ways for future work. Long-term stability and any processes for calibration shift would be characterized by placement studies that last at least one year and cover more than one season. Unit-to-unit error could be measured and production limits set through multi-platform trials with multiple sensor pods working at the same time in the same setting. It would be easier to find any unexpected interactions with vehicle systems, electromagnetic surroundings, or operating processes if integration studies were done on real autonomous vehicle platforms instead of instrumented test cars. As production goes up, economic research that keeps track of real purchase costs would confirm or improve the cost estimates and find ways to cut costs through design for manufacturability. Lastly, basic study into quantum devices that work at room temperature could someday get rid of the need for cold power. This would make the range of possible uses much wider, including smaller cars, missions that last longer, and platforms that do not have a lot of power.

The results of this study have effects on more than just urban travel. They have effects on the whole field of quantum technology. The successful example of active environmental adjustment through quantum-classical co-design goes against the common belief that to keep working properly, quantum devices need to be kept away from their surroundings. This way of thinking comes from basic physics studies where isolation is both possible and important. It has kept people from using quantum sensors in real-world situations where separation is fundamentally limited. The results show a different way of thinking: instead of trying to make the lab conditions work in the real world, quantum systems can be made to sense their surroundings and change to fit them, using the same quantum resources that make them sensitive to allow them to make them compensate. This idea could be used for other quantum technologies too, like quantum communication (which can adapt to changes in the atmosphere or fibers), quantum computing (which can adapt to noise patterns specific to hardware), and quantum time (which can adapt to changes in gravity). Even though this adaptable feedback design was made for guidance, it follows some general rules that may be useful in all areas of quantum technology.

The data from the interviews with stakeholders give more information than just the numbers. Quantum scientists were surprised to learn that it is possible to keep 98.2% of coherence during shaking. This shows that the field's standards have been shaped more by lab examples than by engineering facts. Urban planners are focusing on cost thresholds of $150,000 to $250,000 and municipal engineers are focusing on integration timelines of 72 hours. This shows how important it is to deal with non-technical adoption barriers early in the technology development process instead of putting them off until technical performance is established. The framework's flexible structure, well-documented APIs, and ability to work with current traffic management systems show that it was designed with institutional compatibility and technical performance in mind, which was supported by the replies from stakeholders. It is clear what this means for the future growth of quantum technology: to go from the lab to the real world, you need to optimize the technical, economic, and political aspects all at the same time, instead of focusing on each one separately.

It is important to talk about how this work relates to the DARPA QuANET and EU Quantum Flagship projects because it helps put the framework in the bigger picture of quantum technology. DARPA QuANET has set technology goals for quantum guidance for 2025, such as sub-10-meter accuracy with paths to sub-meter accuracy, power under fifty watts, and 24-hour operation. The system does better than all these goals, with accuracy of 0.15 meters, forty-five watts of power, and 95.3% stability over 24 hours. This match is not a coincidence; it shows that the text used DARPA goals as design guidelines from the very beginning of the project. The SAE Level 4 compliance and performance in civilian urban settings of the framework support the EU Quantum Flagship's focus on social uses, such as self-driving cars and protecting key infrastructure. In direct response to the Flagship's request for technology change roadmaps, the policy-ready framework is made up of cost analyses and approval paths. This work will affect both the U.S. and European strategies for quantum technology because it aligns multiple programs. It may also make it easier for countries to work together on standards, testing procedures, and legal frameworks.

The statistical methods used, especially the stacked ANOVA with interaction terms and the Monte Carlo sensitivity analysis, are more rigorous than what is usually done in the guidance literature. A lot of work on guidance report gains in mean accuracy without explaining variations, relationship effects, or statistical significance well enough [2, 3, 5]. This paper gives a complete statistical analysis that includes power analysis (more than 0.99 for main effects), effect size reporting ($\eta^2 = 0.38$, Cohen's d up to 2.34), and distributional characterization (kurtosis, skewness, and tail percentiles). It can be used as a model for comparing navigation technologies in the future. The sensitivity analysis using Sobol indices found that vibration coupling (34%) and EMI susceptibility (22%) are the main sources of uncertainty. This goes beyond just characterizing the problem and helps set goals for future research and development. This level of scientific rigor is necessary for claims of quantum benefit to be believed, especially since literature on quantum technology has a history of making false claims.

The proof of repeatability, with 0.16 ± 0.04 m RMSE compared to 0.15 ± 0.03 m in the original experiment and a correlation value of 0.94, answers an important question in experimental science in general and in quantum technology in particular. The replication problem in the biological and psychological sciences has made people more aware of the fact that many published results may just be the result of odd lab conditions or flexible analysis and not stable phenomena [23]. The important level of repeatability here, which was reached by using thorough rules, separate hardware, and different study teams, gives us faith that the performance the study saw is due to the framework itself and not just the application. Putting raw data, analysis scripts, and interview notes in a public source makes it easier to do the same thing again and lets other people check it.

Using both qualitative talks with stakeholders and quantitative characterization of performance is a new way of doing things that should be used more often in study on technology development. Adoption needs that quantitative measures alone could not show were found by the theme analysis. These needs include institutional belief in modern technology, fitting with current buying processes, availability of upkeep and support, and agreement with political goals. Even though these things are not shown in RMSE or power consumption numbers, they may affect whether technology is used, even if it works technically. As important for adoption as the technical validation is the fact that the cost of $210,000 is within the acceptable range set by urban planners ($150,000–$250,000). Also, the framework's 72-hour integration timeline directly addresses the municipal engineer's concern about deployment disruption. Similar early partner involvement should be thought about for future technology development projects to make sure that technical specs match the needs of real-world usage.

In the end, the discussion put the experimental results into the context of the translational gaps found in the literature. It also looked at the mechanisms that led to the performance that was seen, as well as the limitations and boundary conditions. Finally, it talked about what these findings mean for future research and the transition to modern technologies. The framework's accuracy of 0.15 meters, dependability of 95.3%, and HRL-7 approval all show that quantum guidance has hit a level of maturity that makes it technically and financially possible to use in smart cities. These designs—adaptive feedback architecture, dynamic mixing, and thorough evaluation methodology—can be used as a model for future quantum technology development that puts environmental stability and institutional compatibility ahead of raw performance. The problems found, like the possibility of magnetic anomalies and the need for cold power, make it clear what needs to be done in future study without taking away from what has already been done. With this work, the quantum navigation field goes from asking "if quantum advantage is possible" to asking "how to deploy quantum navigation systems at scale." This is a substantial change in the research goal that will shape the next stage of growth.

6. Conclusions

This study showed that quantum-enhanced navigation, when used with the suggested adaptive QAOA-AI hybrid architecture, works well enough to be used in urban areas that do not have GNSS. This closes the gap between lab demonstrations and real-world needs. The framework's 0.15-meter root mean square positioning accuracy, 95.3% operational reliability during sustained signal denial, and 82% reduction in drift compared to LiDAR-SLAM baselines show that quantum sensing technologies have grown up enough to be useful in smart city infrastructure, not just for basic physics research. These accomplishments are not small steps forward; instead, they mark a meaningful change in the conditions under which self-driving cars can operate, making it possible for SAE Level 4 features to be used in places where traditional guidance systems simply cannot keep up with performance needs.

Not only does this work show better performance, but it also shows that environmental decoherence, which has been thought to be the biggest problem with using quantum sensors outside of labs for a long time, can be actively fixed through quantum-classical co-design instead of just being tolerated or left alone. The adaptive feedback architecture, which includes a PPO agent that constantly updates the QAOA parameters based on urban telemetry, kept 98.2% of its coherence under ISO 16750-3 vibration profiles and lost less than 1% of its performance under IEC 61000-4-3 electromagnetic interference.

This new way of looking at external noise as a limit on dynamic optimization instead of a source of static errors makes it easier to use quantum technologies in a wide range of demanding settings, from space systems to industrial sites to platforms in the air. The reinforcement learning agent changed the dynamic mixer Hamiltonian to steer quantum exploration toward solution regions that make sense in the real world. This led to 35% higher qubit efficiency than quantum annealing baselines, showing that adaptive algorithm design can help make up for hardware limitations in the NISQ era.

The rigorous testing method includes digital twin simulations (10,000 Monte Carlo runs per scenario), hardware-in-the-loop execution on superconducting quantum processors (IBM Quantum Falcon family), and field trials covering 47.3 kilometres of GNSS-denied routes in three cities. This sets a new standard for accuracy in quantum navigation research. The statistical analysis, which included multifactorial ANOVA ($F(3,1196) = 87.3$, $p < 0.001$, $\eta^2 = 0.38$), effect size quantification (Cohen's d up to 2.34), distributional characterization (kurtosis = 4.2, skewness = 0.1), and sensitivity analysis via Sobol indices, goes above and beyond what is usually done in the navigation literature. It gives confidence that the improvements seen are both statistically significant and useful in real life.

The evidence of repeatability, with 0.16 ± 0.04 m RMSE achieved by separate replication compared to the original 0.15 ± 0.03 m (correlation 0.94), allays worries

about lab effects that are unique and supports the results being applicable to other situations. Archiving raw data, analysis scripts, and interview records for everyone to see makes it possible for independent proof and makes openness a central concept of the work.

The HRL-7 certification is the first time in the peer-reviewed literature that a quantum navigation system has been evaluated against all the criteria needed for cities to adopt it. These include technical performance (0.15 m accuracy, 45 W power, twenty ms latency), economic viability ($210,000 per unit at volume one hundred, sensitivity to $120,000 at volume 1,000), and institutional compatibility (SAE Level 4 compliance, DARPA QuANET alignment, EU Quantum Flagship congruence).

The talks with twenty-four stakeholders confirmed that these criteria are in line with what is needed for adoption. The cost falls within the $150,000 to $250,000 range, which is acceptable to urban planners, and the 72-hour integration timeline meets the standards of city engineers. The framework can be easily integrated with current traffic management systems because it uses detailed APIs and a flexible design. This means that installers do not need to have special knowledge in quantum computing to do so. This removes a non-technical barrier that has traditionally kept people from using advanced guidance technologies.

While proving general stability, the environmental assessment also found specific weaknesses that set the working range and guide future development. The 8–12% drop in performance in steel-rich environments is caused by magnetic field distortions that affect NV centre measurements. To fix this, the system needs to be pre-calibrated against local magnetic anomaly maps, which is a manageable process like the pre-surveying that is already done for lane-level digital maps.

The cryogenic power consumption of 32.1 watts (71% of total) shows that room-temperature quantum sensors should be a top priority for future research. Getting rid of the cryocooler would make the applications much wider, including smaller vehicles, missions that last longer, and platforms that do not have a lot of power. The Pareto analysis that shows declining returns beyond p = 8 QAOA levels (0.7% accuracy gain for 60% power increase) helps engineers figure out how to make the system work better. It shows that the way it is now is at the bottom of the power-accuracy curve.

This work has implications for theory that go beyond guidance and cover a wider area of quantum technology. The successful example of active environmental compensation goes against the common belief that quantum sensors need to be cut off from their surroundings to stay coherent. This belief comes from basic physics tests where separation is both possible and necessary. The results point to a unique way of thinking: quantum systems can be made to sense their surroundings and change to fit them, using the same quantum resources that make them sensitive to allow them to also make them compensate.

This idea could be used in other areas as well, like quantum communication (adjusting to changes in the atmosphere or fibres), quantum computing (adjusting to noise patterns unique to hardware), and quantum time (adjusting to changes in gravity). The adaptive feedback system that was created here was specifically designed for guidance, but it follows general rules that may apply to all quantum technologies. This could speed up the process of putting many quantum technologies into use from the lab to the real world.

There are some problems with this study that should help guide future research. Even though the field tests were very thorough, they were only done with one sensor pod for a brief time. To find out about long-term stability, calibration drift, and unit-to-unit variability, the sensors would have to be used for longer periods of time, across more than one season, and on more than one platform. Even though the magnetic anomaly risk is doable, it means that deploying the system across the country or the world would take a lot of work ahead of time to do a study. However, the plans that are made could be used by different ships. The full cost analysis is based on current prices for low-volume production. Actual costs at scale may be different from estimates because of learning curves in manufacturing, changes in the supply chain, or modern technologies that compete with the one being used.

Extreme conditions that were not tried (for example, vibrations at the level of an earthquake or electromagnetic pulse events) have not been described yet, and they may show failure modes that were not seen in this program. These limits set a clear research agenda: longer deployment studies, trials on multiple platforms, merging with autonomous car platforms, economic research that tracks real buying costs, and basic research into quantum devices that work at room temperature.

Because this work is connected to national and foreign quantum technology projects, it has the potential to change policy and funding objectives. The system goes beyond all of DARPA's QuANET 2025 technology goals (0.15 m vs. <10 m accuracy, 45 W vs. <50 W power, 95.3% vs. >90% dependability), showing that the United States' investment in quantum sensing is paying off with usable capabilities. The connection with EU Quantum Flagship social application priorities, such as protecting key infrastructure and self-driving cars, helps Europe reach its goals for technology freedom and industrial competitiveness.

The HRL-7 license gives a shared language for talking about development in scientific, practical, and administrative areas. This makes it easier for people from different countries to work together on standards, testing procedures, and rules and regulations. Cost analysis and buying paths give government and business leaders the economic information they need to plan budgets and make business cases.

The new ways of doing things described in this paper—the adaptive feedback architecture, the dynamic mixer Hamiltonian, the comprehensive six-dimensional

evaluation grid, and the HRL assessment—can be used as a model for future work on quantum technology that puts environmental robustness and institutional compatibility ahead of raw performance. When you combine quantitative performance evaluation with qualitative stakeholder involvement, you can see that to successfully translate technology, you need to make improvements in all three areas at the same time, rather than focusing on each one separately. The use of statistical accuracy, showing that results can be repeated, and open data sets standards for the quality of proof that should be used more often in quantum technology research, especially since claims have been overstated in the past.

In conclusion, this study has shown that quantum-enhanced urban guidance is no longer just a cool idea for the lab. It is now a real technology that can be used in the real world, with well-known pros and cons and effortless ways to put it into practice. The QAOA-AI system gets 0.15-meter accuracy in places where GNSS is not available, 95.3% reliability when there is a sustained loss of signal, and HRL-7 approval across all policy-relevant dimensions. These accomplishments fill in the gap in translational research found in the literature. They give city governments, companies that make self-driving cars, and defence contractors the empirically grounded, economically validated data they need to make choices about adoption.

With this framework, the quantum navigation field goes from asking "if quantum advantage is possible" to asking "how to deploy quantum navigation systems at scale." This is a substantial change in the research goal that will shape the next stage of growth. As quantum hardware keeps getting better, coherence times get longer, and costs go down, the ideas shown here will make systems that are even more useful. Eventually, the dream of quantum-enhanced smart cities where self-driving cars work safely and reliably even when GPS signals are not working will come true.

Acknowledgments.

The author would like to thank all those involved in the work who made it possible to achieve the objectives of the research study.

References

[1] X. Song, H. He, Z. Zeng, Y. Peng, and W. Hong, "Performance evaluation method for GNSS/INS integrated navigation system," *Sixth International Conference on Geoscience and Remote Sensing Mapping (GRSM 2024),* vol. 13506, pp. 198–210, 2025, doi: https://doi.org/10.1117/12.3057564.

[2] P. Wang, Y. Gao, Q. Zhao, Y. Wang, F. Zhou, and D. Zhang, "An Enhanced, Real-Time, Low-Cost GNSS/INS Integrated Navigation Algorithm and Its Platform Design," *Sensors (Basel, Switzerland),* vol. 25, no. 7, p. 2119, 2025, doi: https://doi.org/10.3390/s25072119.

[3] O. Elmaghraby, P. R. M. Araujo, S. I. K. Abdelaziz, and A. Noureldin, "Comparative Study of Traditional and Deep Learning Feature Detectors and Matchers for Land Vehicle Monocular Visual Odometry," *2025 IEEE International systems Conference (SysCon),* pp. 1–8, 7–10 April 2025 2025, doi: https://doi.org/10.1109/SysCon64521.2025.11014855.

[4] J. Shang, Y. Liu, Y. Xu, J. Xiao, and D. Ma, "Robust Global Localization for Urban Autonomous Vehicles via 3D Geometric-Enhanced Visual Place Recognition," *IEEE Transactions on Intelligent Transportation Systems,* pp. 1–15, 2025, doi: https://doi.org/10.1109/TITS.2025.3590549.

[5] C. Ban, L. Wang, R. Chi, T. Su, and Y. Ma, "A Camera-LiDAR-IMU fusion method for real-time extraction of navigation line between maize field rows," *Computers and Electronics in Agriculture,* vol. 223, p. 109114, 2024/08/01/ 2024, doi: https://doi.org/10.1016/j.compag.2024.109114.

[6] M. Stefanoni, I. Kovács, P. Sarcevic, and Á. Odry, "A Survey on the Main Techniques Adopted in Indoor and Outdoor Localization," *Electronics,* vol. 14, no. 10, p. 2069, 2025, doi: https://doi.org/10.3390/electronics14102069.

[7] J. Bernard *et al.*, "Atom interferometry using σ+-σ− Raman transitions between| F= 1, m F=∓ 1⟩ and| F= 2, m F=±1⟩," *Physical Review A,* vol. 105, no. 3, p. 033318, 2022, doi: https://doi.org/10.1103/PhysRevA.105.033318.

[8] X.-L. Chen *et al.*, "Robust large-momentum-transfer atom interferometry with Raman adiabatic rapid passage," *Optics Communications,* vol. 573, p. 131005, 2024/12/15/ 2024, doi: https://doi.org/10.1016/j.optcom.2024.131005.

[9] L. Rossi, M. Reguzzoni, Ö. Koç, G. Rosi, and F. Migliaccio, "Assessment of gravity field recovery from a quantum satellite mission with atomic clocks and cold atom gradiometers," *Quantum Science and Technology,* vol. 8, no. 1, p. 014009, 2022, doi: https://doi.org/10.1088/2058-9565/aca8cc.

[10] H. P. Paudel, G. R. Lander, S. E. Crawford, and Y. Duan, "Sensing at the Nanoscale Using Nitrogen-Vacancy Centers in Diamond: A Model for a Quantum Pressure Sensor," *Nanomaterials,* vol. 14, no. 8, p. 675, 2024, doi: https://doi.org/10.3390/nano14080675.

[11] R. Katsumi, K. Takada, F. Jelezko, and T. Yatsui, "Recent progress in hybrid diamond photonics for quantum information processing and sensing," *Communications Engineering,* vol. 4, no. 1, p. 85, 2025/05/08 2025, doi: https://doi.org/10.1038/s44172-025-00398-2.

[12] Z. K. Qiao *et al.*, "Feasibility Study on Using Atomic Gravimeters for Detecting Urban Underground Spaces," *IEEE Transactions on Instrumentation and Measurement,* vol. 74, pp. 1–11, 2025, doi: https://doi.org/10.1109/TIM.2025.3565251.

[13] B. Stray *et al.*, "Quantum gravity gradiometry for future mass change science," *EPJ Quantum Technology,* vol. 12, no. 1, p. 35, 2025, doi: https://doi.org/10.1140/epjqt/s40507-025-00338-1.

[14] J. Chapman, Z. C. Seskir, S. Belguith, and T. Tryfonas, "Quantum Sensing Technologies and Smart Cities: Opportunities and Challenges," *2024 IEEE International Smart Cities Conference (ISC2),* pp. 1–6, 29 Oct.–1 Nov. 2024 2024, doi: https://doi.org/10.1109/ISC260477.2024.11004264.

[15] R. Fakhimi and H. Validi, "Quantum Approximate Optimization Algorithm (QAOA)," *Encyclopedia of Optimization,* pp. 1–7, 2020, doi: https://doi.org/10.1007/978-3-030-54621-2_854-1.

[16] D. Volpe, G. Orlandi, and G. Turvani, "Improving the Solving of Optimization Problems: A Comprehensive Review of Quantum Approaches," *Quantum Reports,* vol. 7, no. 1, p. 3, 2025, doi: https://doi.org/10.3390/quantum7010003.

[17] M. W. Geda and Y. M. Tang, "Adaptive hybrid quantum-classical computing framework for deep space exploration mission applications," *Journal of Industrial Information Integration,* vol. 44, p. 100803, 2025/03/01/ 2025, doi: https://doi.org/10.1016/j.jii.2025.100803.

[18] J. Zhang and J. Zhang, "Classical Dance-Metaheuristic: a metaheuristic optimization algorithm inspired by classical dance," *Control Systems and Optimization Letters,* vol. 3, no. 2, pp. 165–173, 2025, doi: https://doi.org/10.59247/csol.v3i2.206.

[19] R. Ghanbarzadeh and S. Mirjalili, "A structured review of large language models in metaheuristic optimisation," *Decision Analytics Journal,* vol. 15, p. 100587, 2025/06/01/ 2025, doi: https://doi.org/10.1016/j.dajour.2025.100587.

[20] T. Stuyver and C. W. Coley, "Quantum chemistry-augmented neural networks for reactivity prediction: Performance, generalizability, and explainability," *The Journal of Chemical Physics,* vol. 156, no. 8, 2022, doi: https://doi.org/10.1063/5.0079574.

[21] M. Chiofalo, "Quantum Toolbox for Neurobiology Sensory Systems," *Journal of Physics: Conference Series,* vol. 2948, no. 1, p. 012015, 2025, doi: https://doi.org/10.1088/1742-6596/2948/1/012015.

[22] M. AbuGhanem, "IBM quantum computers: evolution, performance, and future directions," *The Journal of Supercomputing,* vol. 81, no. 5, p. 687, 2025/04/01 2025, doi: https://doi.org/10.1007/s11227-025-07047-7.

[23] V. Braun and V. Clarke, "Conceptual and design thinking for thematic analysis," *Qualitative psychology,* vol. 9, no. 1, p. 3, 2022, doi: https://psycnet.apa.org/doi/10.1037/qup0000196.

[24] A. Rajak, S. Suzuki, A. Dutta, and B. K. Chakrabarti, "Quantum annealing: An overview," *Philosophical Transactions of the Royal Society A,* vol. 381, no. 2241, p. 20210417, 2023, doi: https://doi.org/10.1098/rsta.2021.0417.